Phylogeography and Phylodemography of Two Peat Mosses, *Sphagnum fimbriatum* and *S. squarrosum* in Europe

Dissertation

Zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Péter Szövényi

aus

Ungarn

Promotionskomitee:

Prof. Dr. Peter Linder

Dr. Edwin Urmi

Prof. Dr. J. Jakob Schneller

Zürich 2007

Acknowledgements

First of all, I would like to express my gratitude to my three primary supervisors, Dr. Edwin Urmi, Prof. Dr. Jakob Schneller (Institute of Systematic Botany, University of Zurich) and Dr. Zoltán Tóth (Cryptogam Laboratory, Department of Plant Taxonomy and Ecology, Eötvös Loránd University, Hungary) and to Prof. Dr. Peter Linder who made it possible to conduct my PhD in frame of a collaborative research project at the Institute of Systematic Botany. I especially enjoyed Jakob Schneller's unexpected questions and ideas concerning phylogeography, demography and population biology of bryophytes; Edwin Urmi's expertise in taxonomic and distribution-related questions and his scepticism in inferring species history and demography from genetic data; and Zoltán Tóth's ideas about the geographic distribution and population dynamics of the species, his flexibility in organizing, his inexhaustible support in teaching and travelling. Their help was invaluable to continuously improve and finish my PhD in due time. They were more than formal supervisors of my thesis, their office was always open for any kind of discussion.

I very much appreciate the tolerance, patience and help of my wife, Zsófia Hock, during the period of my PhD, including application, paper and thesis preparation, collecting trips and endless labwork. It would not have been possible to finish this work so quickly without her help and continuous full time support.

I would also render my thanks to numerous European collaborators who helped us in several ways. Special thanks go to Adam Hölzer (Germany), Alain Vanderpoorten (Belgium), Ana Seneca (Portugal), Andrea Fodor (Hungary), Andrej Martinčič (Slovenia), Anna Ganeva (Bulgaria), Baiba Bambe (Lettland), Beáta Papp (Hungary), Beatriz Itten (Swizerland), Blanka Shaw (Bohemia), Christian Dolnik (Germany), Claes Pearsson (Sweden), Erzsébet Szurdoki (Hungary), Harald Zechmeister (Austria), Ildikó László (Hungary), Irina Goldberg (Denmark), José Pujos (Andorra), Jürgen Brandt (Germany), Kai Vellak (Estonia), Liudmila Abramova (Russia), Markus Preussing (Germany), Marta Infante (Spain), Michael Ignatov (Russia), Monika Koperski (Germany), Nele Ingerpuu (Estonia), Nils Cronberg (Sweden), Patxi Heras (Spain), Philippe de Zuttere (France), Pierre Goubet and Christine (France), Risto Virtanen

(Finland), Robert Krisai (Austria), Ron Porley (UK), Rudolf Soltes (Slovakia), Sam Bosanquet (UK), Sebastian Sundberg (Sweden), the late Sigurd Såstad (Norway), Sture Westerberg (Sweden), Zoltán Illyés (Hungary).

Both the Swiss and Hungarian institutes provided me a highly stimulating research environment. Colleagues and friends from both institutes filled my working days with happiness, especially Alessia Guggisberg and Guilhem Mansion, Alex Bernhard, Anita Lendel, Beatriz Itten, Brigitte Marazzi, Burgi Liebst, Chloé Galley, Corinne Furrer and Frank Rutschmann, Cyril Guibert, Erzsébet Szurdoki and Péter Ódor, Evelin Pfeifer, Gabriele Salvo, Heike Hofmann, Josefine Jacksch, Julien Bachelier, Marlon Machado, Merran Matthews, Massimo Mastracci, Monica Quast and Andre Simoes, Niklaus Müller, Norbert Schnyder, Rolf Rutishauser, Urs Landergott, Yannick Städler and more. Like scientific researchers in general, I was also relying on the support of external funding. Without the 33 months long support of the “Forschungskredit der Universität Zürich” it would not have been possible to conduct this research. Travel grant of the SANW (Schweizer Akademie der Naturwissenschaften) is also greatly acknowledged.

Table of Contents

Publication list

The publications listed below were born during my doctoral stay at the Insitute of Systematic Botany between 2004 and 2006. Earlier papers are not listed here. The first three publications include results of earlier studies, but the manuscript preparation was mainly carried out at the Institute of Systematic Botany. Publications are either accepted and in press or are in the stage of review. Publications 4, 5 and 6 describe the results of my PhD and are included in the thesis.

1. **Szövényi P**, Hock Zs, Tóth Z (2004) Phorophyte preferences of epiphytic bryophytes in a stream valley in the Carpathian Basin. *Journal of Bryology* **26**, 137-146.

2. Hock Zs, **Szövényi P**, Tóth Z (2004) Seasonal variation in the diaspore bank of bryophytes in open dolomite rock grasslands. *Journal of Bryology* **26**, 285-292.

3. Hock Zs, **Szövényi P**, Tóth Z (2006) Seasonal variation in the spore bank of ferns in grasslands on dolomite rock. *Plant Ecology*, in press, doi: 10.1007/s11258-006-9142-3.

4. **Szövényi P**, Hock Zs, Urmi E, Schneller JJ (2006) New primers for amplifying the *GapC* gene of bryophytes and its utility in intraspecific phylogenies in the genus *Sphagnum*. *Lindbergia*, in press, L957.

5. **Szövényi P**, Hock Zs, Urmi E, Schneller JJ (2006) Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *S. squarrosum* (Bryophyta, Sphagnopsida) in Europe. *New Phytologist*, in press, doi: 10.1111/j.1469-8137.2006.01870.x.

Articles currently in review

6. **Szövényi P**, Hock Zs, Schneller JJ (2006) Contrasting patterns of multilocus genetic structure of two peat moss species in Europe: demography, mating system difference or historical factors? in review, *Molecular Ecology*.

7. Arroyo K, **Szövényi P**, Guggisberg A, Conti E (2006) Effects of Pleistocene Glaciations and Life History on the Genetic Diversity of *Saxifraga florulenta* (Saxifragaceae), a Rare Endemic of the Maritime Alps. in review, *Annals of Botany*.

8. Hock Zs, **Szövényi P**, Tóth Z (2006) Relations between bryophyte diaspore bank and standing vegetation in adjacent open and closed grasslands on dolomite rock. in review, *Journal of Vegetation Science*.

General introduction

Phylogeography in brief, is the joint study of the genealogical relationships and geographic distributions of organism lineages usually relying on uniparentally inherited molecular markers (Avise 2000). During the last decade, phylogeography has emerged as a very active field of experimental research. Phylogeographic investigations on plants and animals started almost simultaneously, but phylogeographic appropriateness of certain parts of the mitochondrial genome (high mutation rates, uniparental inheritance and ease of use) have resulted in a steep increase of investigations on animals (Avise 2000).

In contrast to animals, mitochondrial genome of plants is generally inappropriate for phylogeographic investigations (but see for instance Godbout et al. 2005) due to its low mutation rate and relatively low structural conservation (Wolfe *et al.* 1987; Hewitt 2000). By contrast, the chloroplast genome turned out to be a promising marker for phylogeographic investigations. However, owing to its low mutation rate, large parts of the genome need to be investigated to achieve the appropriate resolution (Soltis *et al.* 1997). Consequently, phylogeographic studies on plants were hindered by the lack of easily accessible markers with appropriate resolution. In the last decade, advancement in PCR technology, automated sequencing and development of new molecular markers and methods (Sequence Characterised Amplified Regions, low- and single-copy nuclear genes, nuclear/chloroplast microsatellites and AFLPs) have led to a rapid increase in the number of plant phylogeographic studies. In the last few years, European phylogeography of several tree and herb species has been investigated in details and there are an increasing number of recent publications covering the worldwide distribution of species (among others Petit *et al*. 2003; Stehlik 2003; Lascoux *et al*. 2004; Alsos *et al.* 2005; McDaniel & Shaw 2005; Skrede *et al.* 2006).

The ample amount of phylogeographic investigations lead to the recognition of general patterns in the European phylogeography of several plant and animal species (Taberlet *et al.* 1998; Hewitt 2000; Lascoux *et al.* 2004; Schönswetter *et al.* 2005). The similar geographic organisation of intraspecific genetic variability presumes that, at least partly, species-independent external impact has shaped genetic variability of species in Europe (Hewitt 2004). Investigations led to the conclusion that in case of several plant

and animal species, genetically well-separated European clades emerged due to isolation and reduced effective population sizes during the Quaternary glaciations (Hewitt 1996; Lascoux *et al.* 2004). After the retreat of the glaciers, genetically differentiated lineages of the species rapidly migrated usually northwards from these refugia. This rapid migration, involving frequent founder events, usually led to a decreasing trend in genetic diversity towards the north in species, which survived the last glaciations in small, southern refugia (Hewitt 1999). The main refugia were located on the Balkan and Iberian peninsulas and in Italy (Taberlet *et al.* 1998; Hewitt 2000; Lascoux *et al.* 2004). In contrast, species, which have survived in scattered refugial populations close to the ice sheet, usually do not show a northwardly decreasing trend of genetic diversity, and almost no geographic structure; probably because of larger effective population sizes and effective gene flow among populations (Rendel & Ennos 2002; Palmé *et al.* 2003; Lascoux *et al.* 2004).

The rapid accumulation of phylogeographic data and the increasing number of comparative investigations on flowering plants obviously show that main differences in phylogeographic patterns among species exist (Palmé *et al*. 2003a,b; Heuertz *et al*. 2004, 2006; Schönswetter *et al.* 2006). For instance, comparative studies prove that current distribution of European trees and shrubs is strongly correlated with their present genetic structure, consequently, distribution and location of refugia during the Quaternary glaciations are the most important factors determining the current genetic structure of European tree species (Aguinagalde *et al*. 2005).

The first phylogeographic investigations were generally descriptive ones, and only formulated „ad hoc“ (statistically not tested) explanations underlying the genetic patterns discovered (Avise 2000). Later on, the development of the Nested Clade Analysis (Templeton *et al.* 1995) provided a statistical framework to accept or reject different alternative hypotheses, although it still did not take the random error associated with phylogeographic processes into account (Templeton 2004). Knowles and Maddison (2002) introduced the expression statistical phylogeography, emphasizing that phylogeographic investigations should be able to give statistically supported explanations underlying phylogeographic structures detected.

Responses of species to climatic or other external environmental impacts usually include contractions and expansion of their distributional ranges. Changes in area of occurrence are inherently connected to changes in the number of populations and individuals leading to population demographic changes over time. It may not be obvious at the first glance, but population demography leaves its trace in the genetic structure of populations, influencing the shape of the genealogy (Emerson *et al.* 2001; Hein *et al.* 2004). Phylodemography, as a newly emerging field of evolutionary research, uses this information to link population demographic history, genetic variability and phylogeographic structure (Emerson *et al.* 2001). In order to infer demographic parameters from current level of genetic variability, the theory of the coalescent process is applied (Kingman 1982a,b).

The theory of the coalescent has been described in details by Kingman (1982a,b) and constitutes one of the most rapidly developing field of modern genetics. Briefly, the coalescent process is a statistical model describing the history of a sample of alleles backwards in time. This model makes it possible to draw inferences about different properties of a sample, such as N_e (effective population size), θ (neutral mutation parameter), migration and recombination rates (Hudson 1990). The neutral coalescent assumes that all mutations are selectively neutral, thus there is no relationship among the state of an allele and the process, which has generated it. Consequently, it is possible to separate the genealogical process of alleles from the mutational process, which allows an exact statistical description of the neutral coalescent process (coalescent process in a Wright-Fisher population). Although the neutral coalescent is relatively easy to describe and properties of populations can be calculated directly using this model, more complicated situations comprising population subdivision, migration, demographic change, selection or recombination need to be tested by simulations and expected and observed distributions of properties compared (Hein *et al.* 2004).

The usual coalescent process traces lineages backward in time (Rosenberg & Nordborg 2002). When two alleles of the sample met, they are said to have coalesced. Following this argumentation, the coalescent tree can be divided into successive coalescent intervals when lineages coalesce. The expected time elapsed until each successive coalescence event is an exponentially distributed random variable and is

dependent on the effective size of the population under consideration. This relationship gives the opportunity to estimate the effective size of a population using a sample of alleles (Emerson *et al*. 2001; Hein *et al.* 2004).

In our study, we used bryophytes, more specifically peat mosses (the genus *Sphagnum*), representing an ideal model system for phylogeographic and phylodemographic investigations (Shaw *et al.* 2002) for the following reasons: (1) several species show world-wide or even bipolar distributions, and numerous species pairs with similar distributions but different life-history characteristics exist; (2) their capability for long range dispersal by spores suggests a different geographic distribution of neutral genetic variation than in seed plants; (3) haploidy of the gametophores gives a special opportunity to obtain unambiguous sequence information of single copy regions, without cloning; (4) in spite of their suitability for sequence-level analyses, bryophytes have rarely been used to study phylogeography and/or phylodemography.

Peat mosses are an ancient plant group with a worldwide distribution and numerous endangered species (Daniels & Eddy 1985; Smith 2004). Extensive and diverse distributional areas spanning different continents, several closely related species pairs with different life-history characteristics and rapidly accumulating molecular data make them the ideal model system for evolutionary research in bryophytes (Cronberg 1996; Shaw 2000; Cronberg & Natcheva 2002; Shaw *et al.* 2003; Shaw *et al.* 2005a,b; Flatberg *et al.* 2006). Peat mosses build the main mass of peatland vegetation, and it is estimated that one-third of the global carbon pool is stored in northern-peatlands. As a net sink for atmospheric carbon, their role in regulating global climate is crucial (Gorham 1991).

Two peat mosses with very similar current distributions, ecological requirements and breeding systems, but presumably different historical demography were selected for the investigations in Europe (Daniels & Eddy 1985; Smith 2004). *Sphagnum squarrosum* Crome is a monoecious species, which frequently produces sporophytes throughout Europe and has a nearly constant historical population size according to literature-based data and recent observations. *Sphagnum fimbriatum* Wils. is a monoecious, potentially self-fertilizing species, which frequently produces sporophytes and has been presumably spreading in Europe during the last decades (Szurdoki & Ódor 2004; Sundberg *et al.* 2006).

In the first chapter, we used 2000 bps sequence of the chloroplast genome to describe the phylogeographic pattern of the two species in Europe and to test whether differences in their phylogeographic structure reflect contrasting demographic histories or they are connected to differences in life-history traits (Szövényi *et al.* 2006a). Geographic distribution and total amount of genetic variability of the species differed considerably but the chloroplast markers did not provide sufficient resolution to test demographic hypotheses statistically.

To get sufficient resolution for statistical tests, primers amplifying more variable parts of the peat moss genome have been developed and tested. When chloroplast markers fail to give the appropriate resolution required, nuclear markers – especially due to sequence variability of introns of single-copy nuclear genes – may be the appropriate source of data (Strand *et al.* 1997; Bartish *et al.* 2006). Beside their higher sequence-level variability, detection of sequencing associated errors and the availability of coding and non-coding regions makes them ideal for population genetic analyses (Nielsen 2005). After testing several candidate single-or low-copy nuclear genes, the GapC gene has been selected. The second chapter describes the process of primer development, the structure of the gene in peat mosses and a preliminary analysis of variability (Szövényi *et al.* 2006b).

Inference of historical demography and demographic parameters from the current level of sequence variability relies on several basic assumptions, which have to be met, in order to get reliable estimates (Rosenberg & Nordborg 2002). More specifically, models assume no population structure and selective neutrality of mutations (Hein *et al.* 2004). Peat mosses, like all bryophytes, produce numerous small spores**,** which can travel longer distances by wind. It is hypothesized that gene-flow among local populations can be efficient enough to obscure historical population structure (Sundberg 2005; Korpelainen *et al.* 2005; Sundberg *et al.* 2006). Consequently, in case of frequent spore production, at least neighbouring populations can be treated as one panmictic unit and thus used in subsequent analyses.

The requirement of selective neutrality of mutations is a more challenging criterion (McDaniel & Shaw 2005). In theory, demography and selection may result in the same pattern of molecular variation, e.g. rapid coalescent events following long

branches on the coalescent tree (Hein *et al.* 2004). In the mismatch distribution, long branches lead to an excess of low-frequency polymorphisms (Slatkin & Hudson 1991; Rodgers & Harpending 1992). Consequently, populations recovering from a strong positive selection event (selective sweep) may show the same excess of singleton polymorphisms as populations that have recently experienced a demographic expansion (Hein *et al.* 2004). Although selective forces and demography may produce the same pattern of polymorphisms, selection affects the genome locally, whereas demography extends its effect across the whole genome. Consequently, investigating multiple, unlinked loci can help to separate the influence of selection from demography on nucleotide polymorphisms (Wright & Gaut 2004; Nielsen 2005).

In the third chapter we used a multilocus approach (three nuclear regions, appr. 3000 bps) and several accessions covering the European distribution of both species to draw inferences concerning phylogeography, demography and molecular evolutionary patterns of the two species. Multilocus analyses of sequence data were applied using coalescent, Bayesian and maximum likelihood estimations, to test whether genetic data support the hypothesis of the population expansion of *S. fimbriatum* and a more stable population size in *S. squarrosum*.

The thesis comprises four chapters: three of them have already been published (available as early-online or in press). The three chapters represent three different genetic methods to approach the question of phylogeography and historical demography of two peat mosses. They cover various analytical methods applicable on the level of sequence variability or fragment polymorphisms. I wish the reader an exciting reading and hope that she/he will be as much fascinated by the wonderful world of bryophytes as I am.

References

Aguinagalde I, Hampe A, Mohatny A, Martin JP, Duminil J, Petit RJ (2005) Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography* **32**, 329-339.

Alsos IG, Engelskjon T, Gielly L, Taberlet P, Brochmann C (2005) Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Molecular Ecology*, **14**, 2739-2753.

Avise JC. 2000. *Phylogeography, the history and formation of species*. Cambridge, MA, USA, Harvard University Press.

Bartish IV, Kadereit JW, Comes HP (2006) Late Quaternary history of *Hippophaë rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (*Chsi*) sequences and chloroplast DNA variation. *Molecular Ecology*, **15**, 4065-4085.

Cronberg N (1996) Clonal structure and fertility in a sympatric population of the peat mosses *Sphagnum rubellum* and *Sphagnum capillifolium*. *Canadian Journal of Botany*, **74**, 1375-1385.

Cronberg N, Natcheva R (2002) Hybridization between the peat mosses, *Sphagnum capillifolium* and *S. quinquefarium* (Sphagnaceae, Bryophyta) as inferred by morphological characters and isozyme markers. *Plant Systematics and Evolution*, **234**, 53-70.

Daniels RE, Eddy A (1985) Handbook of European *Sphagna*. Institute of Terrestrial Ecology, UK, Huntingdon.

Emerson BC, Paradis E, Thébaud C (2001) Revealing the demographyc histories of species using DNA sequences. *Trends in Ecology and Evolution*, **16**, 707-716.

Flatberg KI, Thingsgaard K, Såstad SM (2006) Interploidal gene flow and introgression in bryophytes. *Sphagnum girgensohnii* x S. *russowii*, a case of spontaneous neotriploidy. *Journal of Bryology*, **28**, 27-37.

Godbout J, Jaramillo-Correa JP, Beaulieu J, Bousquet J (2005) A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology*, **14**, 3497-3512.

Gorham E (1991) Northern peatlands : role in the carbon cycle and probably responses to climate warming. *Ecological Applications*, **1**, 182-195.

Hein J, Schierup MH, Wiuf C (2004) Gene Genealogies, Variation and Evolution. A Primer in Coalescent Theory. Oxford University Press, Oxford.

Heuertz M, Fineschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG (2004). Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology*, **13**, 3437- 3452.

Heuertz M, Fineschi S, Sebastiani F, Hausman F, Paule L, Vendramin GG (2006) Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. *Molecular Ecology*, **25**, 2131-2140.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.

Hewitt GM (1999) Postglacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **359**, 183-195.

Hudson RR (1990) Gene genealogies and the coalescent process. In: Harvey PH, Partridge L, eds. *Oxford Surveys in Evolutionary Biology. Vol. 7.* Oxford University Press, USA, New York, 1-44.

Kingman JFC (1982a) The coalescent. *Stochastic Processes and Applications*, **13**, 235-248.

Kingman JFC (1982b) On the genealogy of large populations. *Journal of Applied Probability*. **19A**, 27-43.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Korpelainen H, Pohjamo M, Laaka-Lindberg S (2005) How efficiently does bryophyte dispersal lead to gene flow? *Journal of the Hattory Botanical Laboratory*, **97**, 195-207.

Lascoux M, Palmé AE, Cheddadi R, Latta RG (2004) Impact of ice ages on the genetic structure of trees and shrubs. *Philosophycal Transactions of the Royal Society of London, Series B*, **359**, 197-207.

McDaniel SF, Shaw AJ (2005) Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. *Molecular Ecology*, **14**, 1121-1132.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197-218.

Palmé AE, Semerikov V, Lascoux M (2003a) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix cinerea*. *Heredity*, **91**, 465-474.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003b) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology* **12**, 201-212.

Petit RJ, Aguinagalde I, de Bealuiei JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohatny A, Müller-Starck G, Demesure-Musch B, Palmé A, Martin JP, Rendell S, Vendramin GG (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563-1565.

Rendel S, Ennos RA (2002) Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69-78.

Rodgers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, **3**, 380-390.

Schönswetter P, Popp M, Brochmann C (2006) Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species *Minuartia biflora* and *Ranunculus pygmeus*. *Molecular Ecology*, **15**, 709-720.

Scönswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence from glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547-3555.

Shaw AJ (2000) Phylogeny of the Sphagnopsida based on nuclear and chloroplast DNA sequences. *Bryologist*, **103**, 277-306.

Shaw AJ, Cox CJ, Boles SB (2003) Global patterns in peatmoss biodiversity. *Molecular Ecology*, **12**, 2553-2570.

Shaw AJ, Cox CJ, Boles SB (2005b) Phylogeny, species delimitation, and sexual promiscuity in *Sphagnum* section *Acutifolia*. *Systematic Botany*, **30**, 16-33.

Shaw AJ, McDaniel SF, Werner O, Ros RM (2002) Phylogeography and phylodemography. *The Bryologist*, **105**, 373-383.

Shaw AJ, Melosik I, Cox CJ, Boles SB (2005a) Divergent and reticulate evolution in closely related species of *Sphagnum* section *Subsecunda*. *The Bryologist*, **108**, 363-376.

Skrede I, Eidesen PB, Pineiro Portela R, Brochman C (2006) Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Molecular Ecology*, **15**, 1827-1840.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555-562.

Smith AJE (2004) The moss flora of Britain and Ireland. Second edition, Cambridge Univeristy Press, UK: Cambridge.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353-373.

Stehlik I (2003) Resistance or emigration? Response of alpine plants to the ice ages. Taxon, 52, 499-510.

Strand AE, Leebens-Mack J, Milligan BG (1997) Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology*, **6**, 113-118.

Sundberg S (2005) Sporophyte size positively influences short-range spore dispersal in *Sphagnum*, but what happens further away? *Oikos*, **108**, 115-124.

Sundberg S, Hansson J, Rydin H (2006) Colonization of *Sphagnum* on land uplift islands in the Baltic Sea: time, area, distance and life history. *Journal of Biogeography*, **33**, 1479-1491.

Szövényi P, Hock Zs, Urmi E, Schneller JJ (2006a) Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *S. squarrosum* (Bryophyta, Sphagnopsida) in Europe. *New Phytologist*, in press, doi: 10.1111/j.1469-8137.2006.01870.x.

Szövényi P, Hock Zs, Urmi E, Schneller JJ (2006b) New primers for amplifying the *GapC* gene of bryophytes and its utility in intraspecific phylogenies in the genus *Sphagnum*. *Lindbergia*, in press, L957.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789-809.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution ofmitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767-782.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitutuion vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the National Academy of Science USA*, **84**, 9054-9058.

Wright SI, Gaut BS (2004) Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution*, **22**, 506-519.

Chapter I.

Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *S. squarrosum* (Bryophyta, Sphagnopsida) in Europe.

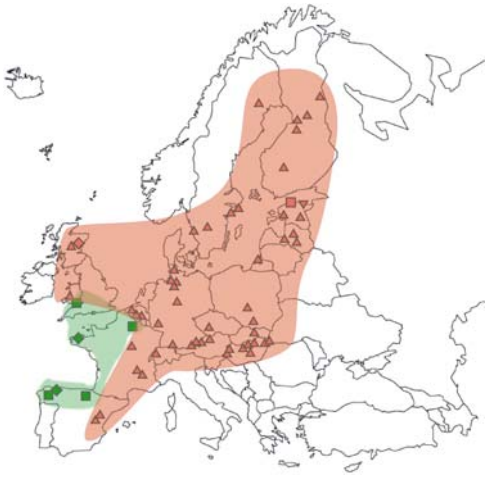
Péter Szövényi[1,2], Zsófia Hock[1,2], Edwin Urmi[1] and Jakob J. Schneller[1]

[1] *Institute of Systematic Botany, University of Zürich, Zollikerstr. 107, CH-8008 Zürich, Switzerland*

[2] *Cryptogam Laboratory, Department of Plant Taxonomy and Ecology, Eötvös Loránd University, Faculty of Science, Pázmány P. sétány 1/C, H-1117 Budapest, Hungary*

Published in

New Phytologist: doi: 10.1111/j.1469-8137.2006.01870.x. 2006



Abstract

The chloroplast phylogeography of two peat mosses *Sphagnum fimbriatum* and *Sphagnum squarrosum* with similar distributions but different life history characteristics was investigated in Europe. Our main aim was to test whether similar distributions reflect similar phylogeographic and phylodemographic processes.

Accessions covering the European distribution of each species were collected and approximately 2000 bps of the chloroplast genome were sequenced. Maximum parsimony, statistical parsimony and phylodemographic analyses were used to address the question of whether these species with similar distributions show evidence of similar phylogeographic and phylodemographic processes.

Chloroplast haplotypes of the currently spreading *Sphagnum fimbriatum* showed strong geographic structure, whereas those of *S. squarrosum*, which has stable historical population sizes, showed only very weak geographic affinity and were widely distributed.

We hypothesize that *S. fimbriatum* survived the last glaciations along the Atlantic coast of Europe, whereas *S. squarrosum* might have had numerous, scattered refugia in Europe. The dominance of one haplotype of *S. fimbriatum* across almost all of Europe suggests rapid colonization after the last glacial maximum. We hypothesize that high colonizing ability is an inherent characteristic of the species and its recent expansion in Europe is a response to climate change.

Keywords: chloroplast DNA, dispersal, glacial refugia, phylodemography, phylogeography, *Sphagnum*.

Introduction

During the last decade, phylogeography has arisen as a promising field of evolutionary research, combining distributional information and phylogenetics to discriminate between past and recurrent gene flow (Templeton, 2004). Phylogeographic investigations of plants are relatively sparse compared to the overwhelming information on animals (Avise, 2000). Studies on non-model plants are generally hampered by the lack of available markers with appropriate resolution. Although the evolutionary rate of the chloroplast genome is relatively slow, several studies using RFLP or sequence data (or both) reported sufficient variation for a phylogeographic time scale (Petit *et al*., 2003; Lascoux *et al*., 2004). The spatial distribution of genetic diversity in European trees and shrubs has been intensively investigated, and general patterns found have been attributed to the effects of the last glaciation (Petit *et al*., 2003).

In general, bryophytes have broad distributional ranges, which can span over continents. Isozymes have revealed that morphological uniformity of world-wide distributed species, in some cases at least, masks complex genetic pattern of intra-specific differentiation (reviewed in Shaw, 2001). Cryptic species have also been found within Europe, including both liverworts and mosses, indicating strong genetic differentiation between populations at a relatively small geographic scale (Odrzykoski & Sweykowski, 1991; Appelgreen & Cronberg, 1999; Shaw, 2001; Sweykowski *et al*., 2005; Fernandez *et al*., 2006). However, several species with distributions spanning several continents show almost no genetic divergence at all (Shaw *et al*., 2003, Korpelainen *et al*., 2004).

Most bryophytes are adapted to long-range dispersal due to several special features, for example desiccation and freezing tolerance, and the production of small, numerous spores. This suggests that partitioning of neutral genetic variation in bryophytes is different from that in seed-dispersed flowering plants (Van der Velde & Bijlsma, 2001; Sundberg, 2005). Like most flowering plant chloroplasts, bryophyte chloroplasts are also assumed to be uniparentally inherited (presumably maternally see Pacak & Szweykowska-Kulinska, 2003; Jankowiak *et al*., 2005). In theory, chloroplast and mitochondrial markers are expected to show a more pronounced genetic structure than nuclear ones because of the reduced effective population size of organellar genomes

and the restricted dispersal of seeds (Avise, 2000). Although dispersal curve of bryophyte spores is strongly leptokurtic (Longton & Schuster 1983), they still can travel to longer distances than seeds and thus considerably raise the degree of gene flow among populations. If elevated gene flow is efficient enough it can totally obliterate historical genetic pattern (Van der Velde *et al*., 2001; Van der Velde & Bijlsma, 2003).

It is evident that species-specific traits influence the spatial distribution of genetic variability. Correlation among different life history parameters and genetic structure has been investigated several times using meta-analyses (Hamrick & Godt, 1996; Aguinagalde *et al*., 2005). Recent investigations on European trees and shrubs revealed that only two out of ten traits showed significant correlation with genetic structure measured by G_{st} and N_{st}, but that the last glacial maximum (LGM) had much more considerable effect on the distribution of chloroplast haplotypes than species specific traits after correcting for pseudoreplication (Aguinagalde *et al*., 2005). Although boreal or temperate species share similar genetic structure, each taxon had its own pattern of recolonization after the LGM. Also taxa with similar habitat preferences and life history traits might show quite different distribution of chloroplast haplotypes (Palmé *et al*., 2003a, 2003b; Heuertz *et al*., 2004).

Hitherto, little is known about the possible refugia of European bryophyte species. The role of the Balkan Mountains, the northern Spain and the southern part of England as refugia are supported by recent investigations (Cronberg, 1998; Natcheva & Cronberg, 2003; Van der Velde & Bijlsma, 2003). However, species with a current circumboreal distribution are likely to have survived the glacial cycles further to the north (Cronberg, 2004).

Among bryophytes, peat mosses have been the most intensively studied group using molecular markers, although Europe wide phylogeographic investigations are scarce and have been conducted only using isoenzymes or fingerprinting methods (Daniels, 1982, 1985a,b; Cronberg, 1996a,b; Stenøien & Såstad, 1999; Såstad *et al*., 2000; Thingsgaard, 2001 among others). Peat mosses represent an appropriate model system for phylogeographic investigations in bryophytes because of their wide distribution ranges spanning several continents, the presence of species pairs with similar

distribution patterns but different life history characteristics, and the availability of easily identifiable fossil material in the form of spores.

The aim of our study was to investigate the chloroplast phylogeographic structure of *S. fimbriatum* Wils. and *S. squarrosum* Crome in Europe. The selected species have similar distributions and reproductive characteristics but presumably different demographic histories (Daniels & Eddy, 1985; Szurdoki & Ódor, 2005). Both have wide distributional ranges spanning North America and Eurasia. They also occur in the Southern Hemisphere; *S. fimbriatum* has been reported from South-America and *S. squarrosum* from New Zealand (Daniels & Eddy, 1985).

Both *Sphagnum* taxa to be investigated are monoecious and frequently sporulating throughout Europe (Wilcox & Andrus, 1987; Cronberg, 1991). They usually occur in mesotrophic to slightly eutrophic habitats, especially at the edge or around bogs, in shaded places. They may also occur along river and pond banks, in fens and in *Salix*, *Betula* or *Alnus* carrs (Daniels & Eddy, 1985). The European distribution centre of *S. squarrosum* lies in the northern part of the continent. Similarly, *S. fimbriatum* is more abundant in the northern part and typically occurs in lowland conditions (Daniels & Eddy, 1985; Schröck & Krisai, 1999; Feldmeyer-Christe *et al*., 2001). Although they share similar ecological requirements, distributions and reproductive characters, their population demographic histories differ considerably. During the last 20 years, *S. fimbriatum* has been reported as spreading in several parts of Europe, while similar trends have never been observed in the case of *S. squarrosum* (Szurdoki & Ódor, 2005). Based on several observations *S. fimbriatum* successfully colonizes bare open soil surface, even in pebbled mine pits (Szurdoki & Ódor, 2004). Under such conditions its colonization ability is evidently greater than that of *S. squarrosum*, however, investigations concerning the mechanisms of this process are lacking (Sundberg & Rydin, 2002).

The first goal of the present study was to provide a detailed picture about the distribution of chloroplast haplotypes in Europe, with special reference to the central and western European populations. As both species have similar European distribution ranges, we tested the hypothesis that similar distributions reflect similar ancient histories. Since the two taxa also show similar reproductive traits, we were also interested in how species-specific colonization ability influences phylogeographic structure. Lastly we

were looking for signs of the different demographic histories in the chloroplast diversity of the species.

Materials and methods

cpDNA analysis

Specimens of *S. fimbriatum* or *S. squarrosum* have been obtained either from European herbaria or collected by the authors (see in Table 4 and 5). We aimed to cover as much of the European range of the species as possible. Each sample contained a group of 3-5 shoots. Freshly collected specimens were dried in silica gel and stored until DNA analysis at 4 °C. All specimens were checked microscopically and vouchers are deposited in the authors' private herbaria. As initial screening (see regions tested below) of within population variability using four accessions per populations gave no polymorphism usually only one individual per locality was analysed.

Total genomic DNA was extracted from air-dried, silica gel-dried or herbarium specimens using the Dneasy Plant Mini Kit (Qiagen, Hombrechtikon, Switzerland) following the manufacturer`s protocol. Initially seven individuals (of each species) were used for screening covering the European distribution range of the species. Several noncoding regions of the chloroplast genome were tested (*trnD*^{guc}-*trnT*^{ggu} spacer, *atpB-rbcL* spacer, *trnT*^{ugu}-*trnL*^{uaa} spacer, *matK* gene, *petL-psbE* spacer, *psbD-trnT*^{ggu} spacer, *rpS12-rpL20* spacer, *rpoB-trnC*^{gca} spacer, *trnS*^{gcu}-*trnG*^{uuc} spacer, *rpL16* intron, *trnL* intron and *trnL-trnF* spacer, primers see in Shaw *et al.*, 2005) using universal primer pairs or primer pairs developed by ourselves to study DNA polymorphism. Among these three showed considerable variability and consistent amplification, and were therefore chosen for further analysis (*trnL*^{uaa}-*trnL*^{uaa}-*trnF*^{gaa}, *trnS*^{gcu}-*trnG*^{uuc} spacer, and *rpL16* intron, hereafter trnLF, trnSG and rpL16 respectively). To improve the success of amplification from herbarium specimens primer sequences were modified or additional internal/external primers designed to conduct semi-nested PCRs. Primer sequences used in this study and their applications are given in Table 1. PCR amplifications were performed using the external primer pairs. In case of herbarium specimens a second PCR was usually done using 1 µl from the 10 x diluted first round product as template with one internal primer and one primer from the first round (see Table 1). The PCR mix contained 1-2 µl of template DNA (3-5 ng/µl), 11.5 µl ddH_2O, 0.4 µl of each primer (10

µM), 3.2 µl of a 1.25 mM dNTP solution in an equimolar ratio, 0.1 µl Taq DNA-polymerase (5u/µl, Sigma), 2 µl of the provided 10 x enzyme buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM $MgCl_2$, 0.01 % gelatin). PCR reactions were run either on a Biometra (T1 or Tgradient, Whatman Biometra, Göttingen, Germany) or on a Techne ptc 412 (Barloworld Scientific Ltd., Stone, UK) thermocyclers. Except for the trnSG spacer the PCR cycling conditions were 94 ºC denaturating temperature for 4 min, then 94 ºC for 1 min, different annealing temperatures (Table 1) for 1 min and 72 ºC extension for 1-1.5 min for 35 cycles with a final extension at 72 ºC for 7 min. In case of the trnSG spacer the two step program of Shaw *et al*. (2005) was used to prevent the coamplification of the *trnS*^{uga}-*trnG*^{ggc} region. The products were checked on 0.8 % agarose gels and cleaned using the GFX PCR and gel band purification kit (Amersham Biosciences, Otelfingen, Switzerland). Approximately 10 ng product was sequenced in 10 µl cycle sequencing reaction with Big dye v3.1 in an ABI prism 3100 (Applied Biosystems, Rotkreuz, Switzerland) genetic analyzer using the original PCR primers. Sequences were checked, corrected if necessary and contigs made using the Sequencher 4.5 software (Gene Code Corporation, Ann Arbor, MI, USA). After assembling the data matrix all polymorphisms were rechecked and false base callings were manually corrected.

Data analysis

Before combining the different chloroplast data sets visual inspection and the incongruent length difference test were applied as implemented in PAUP4.0beta10 (Swofford, 1999) to test the significance of partition homogeneity against random partitions (10000 random replicates were applied). Both inspection and the formal test supported the same partitions for almost all regions.

Table 1 Usage and sequences of primers used in this study. Modified nucleotides and primer names are in bold. 1st and 2nd refer to the first and the second round of a semi nested PCR. In the second round PCR 1 µl 10x diluted first round product was used as template.

CpRegion	Name	Usage	Sequence (5`-3`)	Design based on	Annealing T
trnL^{uaa}-*trnL*^{uaa}-*trnF*^{gaa}	**trnT^{ugu}Fmod.**	amplification	cattacaa**g**tgc**gg**tgctct	Taberlet et al. (1991)	1st 55 °C
	trnF^{gaa}mod.	amplification and sequencing	atttgaactggtgacac**a**ag	Taberlet et al. (1991)	2 nd 55 °C
	5`trnL^{uaa}F (TabC)	amplification and sequencing	cgaaatcggtagacgctacg	Taberlet et al. (1991)	
trnS^{gcu}-*trnG*^{uuc}-*trnG*^{uuc}	3`trnG^{uuc}	amplification	gtagcgggaatcgaacccgcatc	Shaw et al. (2005)	1st 66 °C
	trnS^{gcu}	amplification and sequencing	agatagggattcgaaccctcggt	Shaw et al. (2005)	2nd 66 °C
	5`trnG2S	amplification and sequencing	ttttaccactaaactatacccgc	Shaw et al. (2005)	
rpL16 intron	rps3 5`end	amplification	ccagctcaaacaatttatggagt	*Physcomitrella patens* sequences	1st 55 °C
	rpl16F71mod.	amplification and sequencing	**gtt**atgcttagtgt**ac**gactcgt	Baum et al. (1998)	2nd 57 °C
	rpl16R1516	amplification and sequencing	cccttcattcttcctctatgttg	Baum et al. (1998)	

Based on their congruence, data sets of the three regions were combined to increase resolution. Owing to the relatively low variability, each accession was manually assigned to one of the haplotypes found. After that, a new data matrix was constructed containing only one sequence for each haplotype and was used in the phylogenetic analyses. Indels were recoded as presence-absence and were also used in the analyses. The combined dataset was analysed using maximum parsimony in PAUP4.0beta10 (Swofford, 1999). Due to the low number of haplotypes, exhaustive searches were performed. For *S. fimbriatum* one part of the trnLF region containing a chloroplast microsatellite was excluded (Table 2) because of its ambiguous coding and different evolutionary rate. Support of each branch was estimated using 10 000 bootstrap replicates. Trees were rooted using accessions of *S. palustre*, *S. girgensohnii* and *S. teres* as outgroups.

When sequence divergence is low, conventional phylogenetic methods may perform poorly because of several violations from the assumptions of the methods. Statistical parsimony was used to estimate an unrooted haplotype network and the 95 % plausible set of connections between haplotypes (Templeton *et al*., 1992). Analysis was performed as implemented in the TCS software (Clement *et al*., 2000). In case of *S. fimbriatum* the chloroplast microsatellite of the trnLF region was also included and absolute differences in the number of polyA repeats increased the distance between haplotypes. Nested clades were manually delimited applying the rules reported in Templeton et al. (1987) and Templeton and Sing (1993).

To investigate the amount of chloroplast diversity, two estimators of θ were calculated: nucleotide diversity (π) and Wattersons` θ (Hudson, 1990). Tajima`s D (Tajima, 1989) and the R_2 statistic (Ramos-Onsins & Rosas, 2002) were also calculated to look for signs of a population demographic change. Tajima`s D and R_2 statistic are sensitive to the excess of singleton polymorphisms in the data set and might be indicative for population growth or decline. Values of these statistics were calculated either using SITES (Hey & Wakeley, 1997) or DNAsp v.4.10.4 (Rozas *et al*., 2003).

Results

Molecular variation

In *S. fimbriatum* aligned lengths of the regions used were 530 bps (trnLF), 737 bps (trnSG), and 646 bps (rpL16 intron). In total 9 point mutations and 6 indels were found (the chloroplast microsatellite of the trnLF region excluded) (Table 2). Excluding the accession from Svalbard, only two point mutations and 6 indels were detected, of which both point mutations but only one indel was informative (Table 2). The trnL intron of the trnLF region contained a poly-A repeat, which varied in length between 9-15 bps. Otherwise, in terms of point mutations and indels it showed low variability (1 point mutation and no indel, the poly-A repeat excluded). Compared to the other two regions the trnSG intergenic spacer contained 4 point mutations and 3 indels. The intron of the rpL16 gene showed three point mutations and one, four base long indel.

In *S. squarrosum*, aligned lengths of the regions were 535 bps (trnLF), 782 bps (trnSG) and 688 bps (rpL16 intron). In total 9 point mutations and 3 indels were found (Table 2). In one nucleotide position even three substitutions were detected (trnSG, position 1037). All of the point mutations and one of the indels were informative. The poly-A repeat of the trnL intron showed no variability at all, and was disrupted by several point mutations compared to the sequences of *S. fimbriatum*. In the trnLF region both the trnL intron and the intergenic spacer showed variability, each with one point mutation. In the trnSG spacer three point mutations and two indels were detected. The intron of the rpL16 gene contained three point mutations and one six bps long indel.

Gene diversity (measured as π) was an order of magnitude higher in *S. squarrosum* than in *S. fimbriatum* (Table 3). The two estimators of θ differed considerably showing a deviation from the neutral coalescent process, which appears in the estimated values of Tajima`s D statistic (Tajima, 1989) and in the R_2 (Ramos-Onsins & Rosas, 2002) value as well. Also *S. fimbriatum* showed a more negative D value and a lower R_2 value than *S. squarrosum*. Owing to the low resolution of the data set, none of these statistics were significant.

Table 2 Summary of polymorphic sites in the regions studied. Haplotypes are denoted by letters (see Fig. 1) number of accessions for each haplotypes are given in brackets. Position of polymorphic sites in the aligned data set are indicated at the top of the figure. Deletions are represented by asterisks (*) and those of more than 1 bp are in bold. Variable part of the poly-A repeat in the trnLF region of *S. fimbriatum* is shown in italics. Hyphen (-) indicates no change.

Sphagnum fimbriatum

	intron								spacer																
Haplotypes	*trnL-L-F*									*trnS-G* spacer								*rpL16* intron							
																1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	3	4	4	6	6	8	8	9	0	1	3	3	7	7	7	7	8	
	8	8	8	8	9	9	1	6	8	1	4	1	4	3	6	8	0	4	1	1	1	1	2	
	6	7	8	9	0	1	7	9	5	0	0	1	9	8	2	7	2	6	4	5	6	7	9	
Consensus	***A***	*	*	*	*	*	G	C	*	*	*	*	C	G	T	G	T	G	**A**	**T**	**T**	**G**	A	
A (1)	-	***A***	***A***	***A***	-	-	A	T	-	-	-	-	T	A	C	T	C	-	*	*	*	*	C	
B (2)	-	***A***	***A***	***A***	***A***	-	-	-	-	-	-	-	-	-	C	-	-	A	*	*	*	*	-	
C (4)	-	***A***	***A***	***A***	***A***	***A***	-	-	-	-	-	-	-	-	C	-	-	A	*	*	*	*	-	
D (1)	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
E (59)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
F (1)	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
G (1)	-	-	-	-	-	-	-	-	-	C	C	C	-	-	-	-	-	-	-	-	-	-	-	

Sphagnum squarrosum

	intron	spacer																
Haplotypes	*trnL-L-F*		*trnS-G* spacer					*rpL16* intron										
					1	1	1	1	1	1	1	1	1	1	1	1	1	
	1	4	5	9	0	2	2	4	4	4	4	4	4	5	5	6		
	2	6	8	1	3	6	9	8	8	8	8	9	9	3	8	0		
	5	7	7	8	7	0	9	6	7	8	9	0	1	4	1	2		
Consensus	G	C	*	C	A	C	*	*	*	*	*	*	*	C	C	C		
A (4)	-	-	-	A	G	A	-	**A**	**T**	**T**	**A**	**T**	**T**	A	-	A		
B (4)	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	A		
C (1)	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-		
D (13)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
E (4)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-		
F (3)	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
G (12)	A	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-		
H (1)	A	T	-	-	-	-	G	-	-	-	-	-	-	-	-	-		

Geographic distribution of S. fimbriatum *haplotypes*

Both maximum parsimony and statistical parsimony provided the same cladograms of *S. fimbriatum* haplotypes (Fig. 1). The Svalbardian accession formed one

divergent clade. Rooting the cladogram with outgroup accessions placed the root between the haplotype from Svalbard and the rest of the accessions.

Except for the Svalbardian haplotype, all other accessions grouped into two major haplotype groups showing prominent geographic affinity (clade 2-1 and 2-2). Haplotypes of nested clade 2-1 occur in Northern, Central and partly Southern Europe, whereas haplotypes of clade 2-2 are restricted to the Atlantic coasts of Spain, France and Britain. Clade 2-2 contains two haplotypes with similar frequencies. Despite the low number of mutations both groups are well- or moderately-supported based on bootstrap replicates (the chloroplast microsatellite excluded). By including the more variable chloroplast microsatellite region both resolution and distances within and among clades increased, however, this did not influence the general relationships among haplotypes. All three regions supported similar partitions.

The number of chloroplast haplotypes in non-glaciated and glaciated parts of Europe is similar (3-3), however their spatial distribution was very different. Central and northern Europe are almost dominated by one haplotype (haplotype E, Fig. 1), whereas the Iberian Peninsula, west of France and the south-western part of the British Isles contain a mixture of relatively divergent haplotypes.

Geographic distribution of S. squarrosum *haplotypes*

Maximum parsimony and statistical parsimony supported the same results from which two main clades emerged (Fig. 1). Clade 2-1 contains two divergent haplotypes and is separated by at least 2 to 3 mutations from clade 2-2. Both haplotypes of clade 2-1 are rare compared to the remainder. They do not show any geographic affinity and are separated by several mutations. Rooting the cladogram with outgroup accessions, haplotype A emerged as basal.

The other main clade (clade 2-2) is widely distributed and did not contain missing haplotypes, but only the two dominant haplotypes show a weak western-eastern split contrasting to the prominent geographic affinity of thoses found in *S. fimbriatum*. The number of haplotypes across glaciated vs. non-glaciated parts of Europe was similar. The

three investigated chloroplast regions gave similar information and supported the same partitions of haplotypes resulting in a non significant partition homogeneity test.

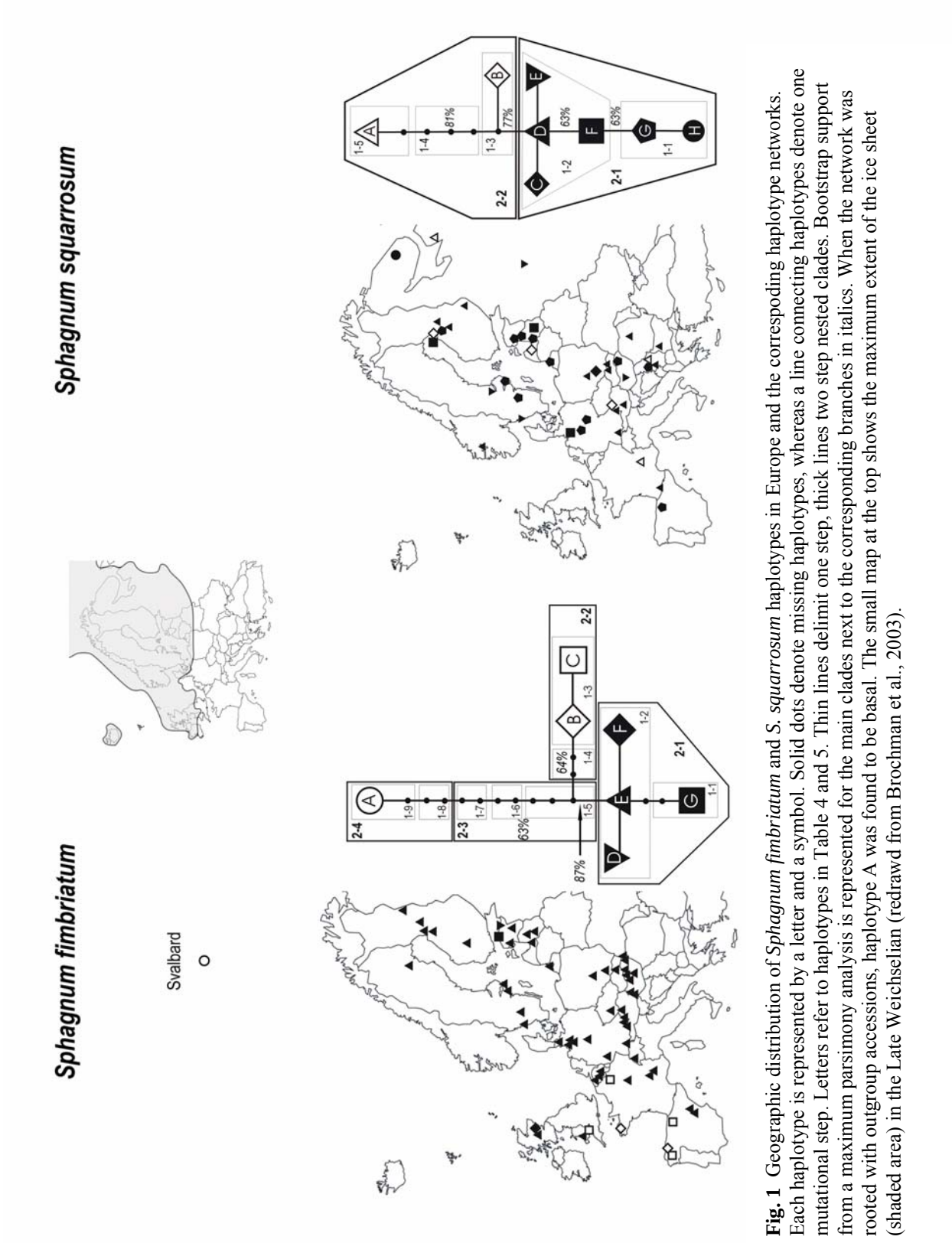


Fig. 1 Geographic distribution of *Sphagnum fimbriatum* and *S. squarrosum* haplotypes in Europe and the correspoding haplotype networks. Each haplotype is represented by a letter and a symbol. Solid dots denote missing haplotypes, whereas a line connecting haplotypes denote one mutational step. Letters refer to haplotypes in Table 4 and 5. Thin lines delimit one step, thick lines two step nested clades. Bootstrap support from a maximum parsimony analysis is represented for the main clades next to the corresponding branches in italics. When the network was rooted with outgroup accessions, haplotype A was found to be basal. The small map at the top shows the maximum extent of the ice sheet (shaded area) in the Late Weichselian (redrawd from Brochman et al., 2003).

Discussion

Geographic distribution of haplotypes and the presence of refugia

S. fimbriatum haplotypes show a distribution with considerable geographic structure (Fig. 1). Haplotypes of nested clade 2-1 occur in Northern, Central and partly Southern Europe, whereas haplotypes of clade 2-2 are restricted to the Atlantic coasts of Spain, France and Britain. Clade 2-2 contains two haplotypes with similar frequencies, however, excluding the chloroplast microsatellite of the trnLF region, would reduce this to one. Given that chloroplast microsatellites evolve 3 or 4 times faster than the average substitution rate of chloroplast genes (Provan *et al*., 2001), it is very likely that length differences among haplotypes arose relatively recently. Considerable genetic differentiation between these two clades suggests isolation of populations.

Data concerning the species composition and extension of *Sphagnum* occurrences during the last glacial maximum (LGM) in Europe are sparse (Gajewski *et al*., 2001). Most of the European tree species survived the last glaciation in Southern refugia like the Iberian Peninsula (Benett *et al*., 1991). Preceding the LGM, parts of the Iberian Peninsula were covered by open woodlands, which might have provided suitable habitats for peat mosses. During the LGM the climate became much more arid but to some extent pine woodlands seem to have existed in sheltered valleys (VanAndel & Tzedakis, 1996). Considerable amounts of peat moss spores have been found in the south-eastern part of Poland and in Northern Spain from about 19 ka before present (Gajewski *et al*., 2001), which might represent possible glacial refugia. Macrofossil remains of *Sphagnum* species have also been found along the Atlantic coast of France, Spain and Britain after the LGM (Forment & Jovet-Ast, 1950; Infante & Heras, 1987; Thingsgaard, 2002). Thingsgaard (2002) speculated that the rare *S. austinii* could have survived the last glaciation along the Atlantic coast of France and Spain. Distribution of the two well-diverged clades of *S. fimbriatum* also supports this hypothesis (Fig. 1). Haplotypes B and C only occur along the Atlantic coasts of Spain, France and Britain. The other clade (2-1) supports the existence of another European refugium.

Looking at the phylogeographic structure of *S. fimbriatum* haplotypes, almost all haplotypes except the rare ones (haplotypes G, D and F) occur on the Iberian Peninsula (Fig. 1). Based on this pattern, a simple scenario can be drawn. Namely a western refugium and an eastern refugium might have existed in the northern part of Spain. Small population sizes and genetic isolation led to the divergence of two groups of haplotypes (clade 2-1 and 2-2). After the retreat of the ice sheet, these haplotypes spread northward and the western group occupied the Atlantic coasts of Britain, France and Spain. The haplotype from Eastern Spain was more successful and recolonized the remaining part of Europe. Based on this scenario, the rare haplotypes (D, F and G) might have evolved relative recently.

Current locations of haplotypes and their putative refugial areas do not necessarily correspond. Establishment and extinction of populations probably occur frequently (Mackay & Talis, 1996; Malmer & Wallen, 1999; Frankl & Schmeidl, 2000), implying a complex history of survival including considerable gene flow between Western Spain, Southern Britain and France. All of these regions might have supported refugial populations during the LGM.

The accession from Svalbard needs to be discussed separately. Strong sequence divergence of this accession implicates a long-standing gene flow barrier that separated it from the rest of the European lineages. Flatberg and Thingsgaard (2003) suggest that colonization of the Svalbard archipelago might have started during warmer periods following the last glacial maximum. They speculate that *Sphagnum* species might have reached the island with the aid of drifting wood or ice from the non-glaciated northern coasts of Eurasian Russia. In our ongoing investigations (unpublished results) this divergent haplotype is found at several locations in Alaska. This, at least in the case of *S. fimbriatum*, contradicts the proposed Siberian origin of the Svalbard lineage. Hybridization among closely related peat moss species appears to be common, thus these individuals also might represent interspecific hybrids (Shaw *et al.* 2005).

Interestingly, *S. squarrosum* haplotypes showed only very limited geographic affinity (Fig. 1). These results are surprising given the strong geographic split found in *S. fimbriatum*. Both haplotypes of clade 2-2 are widespread from France to Archangelsk, but rare haplotypes of clade 2-1 (haplotypes C, E, F, H) show no geographic structure.

The remaining frequent haplotypes (D and G) of the latter clade represent a weak southern-western affinity in their distributions. The two most frequent haplotypes are in the centre of two nested clades (1-2 and 1-1) and represent ancient types according to the rules of the coalescence theory (Crandall & Templeton, 1993). Although this distribution pattern does not give too much information concerning the position of refugial areas, based on the distribution of haplotypes D and G, one might hypothesize the existence of at least a western refugium and an eastern refugium. Based on the previously mentioned investigations on peat mosses (Natcheva & Cronberg, 2003) and *Polytrichum* species (Van der Velde & Bijlsma, 2003), the Balkan Mountains or/and Carpathians might have served as refugia.

Other refugia also seem to have existed as shown by the presence of haplotype B and A, but they cannot be located. In general, one or two substitutions possibly reflect mutations acquired during the LGM (sensu lato 25000-18000 BP) but the six mutations (0.003% difference) between haplotype A and D suggest that their split precedes the last glaciation (0.024-0.116 % divergence/Myr, Hewitt, 2000). In addition, rooting of the cladogram indicates that the first split occurred between haplotype A and the rest of the haplotypes. These facts both show the ancient origin and its marked separation from the other haplotypes. We collected this haplotype only under relatively high altitudinal conditions in Southern and Northern Europe. It might therefore represent an early separated lineage of the species adapted to mountain conditions. This hypothesis is in line with the results found by Cronberg (2004) in *Hylocomium splendens*. He hypothesized that recent populations at higher altitudes represent a lineage that was widely distributed after the LGM on the European tundra. After the climate became warmer, this lineage only survived in alpine habitats. We might hypothesize a similar history for haplotype A. However, its origin is not known yet and only thorough investigations on the world-wide distribution of *S. squarrosum* haplotypes would shed light on this question.

Until now, only two other comprehensive studies investigated the phylogeographic structure of European bryophyte populations across a wide scale. Cronberg (1998) studied the population genetic structure of *S. rubellum* and *S. capillifolium* using isoenzymes. Geographic substructure was found in *S. rubellum*, namely British populations were separated from Scandinavian-German populations.

Higher genetic diversity found in British populations was supposed to indicate refugial populations. Van der Velde and Bijlsma (2003) studied phylogeography of several *Polytrichum* species with a most extensive sampling reaching northern Spain. Based on isozymes, North European populations of *Polytrichum juniperinum* separated well from populations more to the south. Splits within the species were interpreted as colonization of Europe from two different refugia, one in the southern part of Britain and another in Central Europe. These findings are in line with our results supporting Western European refugia in Britain, France or/and Spain.

Haplotype diversity and its implication for population demography

Gene diversity (measured as π) was an order of magnitude higher in *S. squarrosum* than in *S. fimbriatum* (Table 3). This derives from the fact that one haplotype of *S. fimbriatum* was very frequent, while *S. squarrosum* had two main almost equally frequent haplotypes. The two estimatiors of θ differed considerably showing a deviation from the neutral coalescent process. These deviations appear in the estimated values of Tajima`s D statistic (Tajima, 1989) and in the R_2 (Ramos-Onsins & Rosas, 2002) value as well. In the present data set mutations only occurred in non-coding regions. Therefore we might assume that nucleotide changes are not or are only weakly influenced by selective forces and thus evolve under neutral expectations. In such cases, deviations of Tajima`s D statistic from zero and a low R_2 value can be indicative of population growth, decline or bottleneck (Ramos-Onsins & Rosas, 2002). D and R_2 values of both species turned out to be not significant (results not shown) because of the low resolution of the data set. However, *S. fimbriatum* showed a more negative D value and a lower R_2 value than *S. squarrosum*, which indicate an excess of singleton mutations in the data set. This suggests that *S. fimbriatum* went through a severe bottleneck during the last glaciation, which was followed by an exponential population growth and spread. Dominance of almost one haplotype in whole Europe except the Atlantic coast and southern Britain indicates that the spread occurred relatively recently compared to the mutation rate of the regions investigated.

Table 3 Molecular diversity estimates and some species specific lifehistory characteristics of *Sphagnum fimbriatum* and *S. squarrosum* . S: number of segregating sites; π: nucleotide diversity; θw: Watterson`s estimate of the neutral mutation parameter; Tajima`s D(S): estimate based on the number of segregating sites. In case of *S. fimbriatum* the divergent Svalbardian haplotype was excluded. Values in brackets show results including the Svalbardian haplotype.

Species	S	π	θw	Tajima`s D (S)	R_2	Effectiveness of long range dispersal	Desiccation tolerance	Chemical characteristic of substrate for establishment
S squarrosum	8	0,00104	0,00093	0,34792	0,1305	effective	higher	nutrient rich
S. fimbriatum	2 (9)	0.00017 (0.00028)	0.00022 (0.00098)	-0.38631 (-1.87725)	0.0795 (0.0935)	less effective	lower	nutrient poor

In *S. squarrosum*, Tajima`s D statistic was only very slightly negative, and the R_2 value was relatively high (Table 3). Both indicate a relatively constant population size (Ramos-Onsins & Rosas, 2002). The wide distribution and mixing of haplotypes are indicative of long-term persistence of populations as well. Assuming the same mutation rate in both species, *S. squarrosum* may have experienced a less severe bottleneck than *S. fimbriatum* during the last glaciations. A less severe bottleneck probably resulted in the survival of more refugial populations from which haplotypes spread following climate change.

The number of haplotypes observed in the species investigated is very low compared to similar studies on flowering plants (reviewed in Petit *et al*., 2003; McLachlan *et al*., 2005). Taxa with lower number of haplotypes probably underwent more severe bottleneck or have lower mutation rates.

Bryophytes in general show extremely low chloroplast diversity within Europe (Korpelainen *et al*., 2004; Shaw *et al*., 2004; McDaniel & Shaw, 2005). Even using more rapidly evolving regions, only a limited number of haplotypes were detected in European bryophyte taxa (Shaw *et al*., 2003; Shaw *et al*., 2004). *S. fimbriatum* shows this general pattern of lack or limited amount of variance across Europe, with the exception of the samples from Spain. This low level of divergence within Europe is more likely to be a result of a severe bottleneck presumably during the Quaternary glaciations. A

higher number of segregating sites and haplotype diversity in *S. squarrosum* might represent the other extreme of the range, with an only moderate bottleneck.

Species-specific traits vs. phylogeographic structure

Although *S. fimbriatum* and *S. squarrosum* frequently occur together, causes of demographic and phylogeographic differences might lay in the different characteristics of the species. Differences involved in the development of different genetic structure of the species, should be present in at least three stages of their life history: namely establishment preference, effectiveness of dispersal and ecological requirements of the gametophores (Table 3).

The first obvious difference between the two species can be found in their establishment preferences. *S. fimbriatum* is much more successful in colonizing sites with low phosphate content like bare soil or peat, while *S. squarrosum* needs the presence of the appropriate microsites, such as nutrient rich substrate and shading (Sundberg, 2002) for successful establishment. This pattern might have considerably influenced recolonization history of the species after the bottleneck.

The second difference is found in their capability of long distance dispersal. Among the species investigated by Sundberg (1998), *S. squarrosum* had the highest number of spores per capsule and the highest spore output per patch. In addition, the number of spores remaining in the capsules was also low compared to other species. Its capability of long distance dispersal is enhanced by the combination of a large capsule size and relatively small spores (Sundberg, 2005). Although similar data are not yet published concerning *S. fimbriatum*, its smaller capsule size suggests less efficient dispersal because larger capsules are more effective for long distance dispersal than smaller spores (Sundberg & Rydin, 1998; Sundberg, 2005).

Finally, *S. fimbriatum* appears to be less tolerant of desiccation than *S. squarrosum*. Green (1968) showed that all shoots of *S. fimbriatum* died after three days of mild desiccation, while *S. squarrosum* survived much longer. In addition, *S. squarrosum* gametophores appear to be tolerant to several ecological factors: they perform similarly

in rain and ground water and their growth is clearly stimulated by nutrient supply (Kooijman & Bakker, 1995). The climate during the Quaternary glaciations was not only cooler but also much drier than nowadays. Consequently, species with Atlantic affinity and lower desiccation tolerance might have gone through a much stronger bottleneck. Cronberg (1998) investigated two *Sphagnum* species with different climatic and geographic affinities and also detected considerable difference in their genetic structure. He also hypothesized that species with conspicuous Atlantic affinity might have had a more severe bottleneck, which resulted in a marked split in their recent genetic structure. A higher number of nucleotide changes and greater mixing of haplotypes in *S. squarrosum* might have been caused by a moderate bottleneck.

Putting the picture together, it is likely, that species-specific traits can explain, at least part, the differences in phylogeographic structure found in our data. During the Quaternary glaciations *S. fimbriatum* probably went through a more severe bottleneck than *S. squarrosum*. Only few populations with low effective population size of *S. fimbriatum* might have survived. *S. squarrosum* however, could have survived in sheltered valleys in scattered refugia under *Betula* and *Salix* stands more to the north, closer to the margin of the ice sheet. Based on this scenario, we hypothesize that *S. fimbriatum* spread rapidly after the severe bottleneck owing to its good colonizing ability at open soil surfaces. This might have caused a marked genetic and geographic structure. In contrast, *S. squarrosum* might also have spread quite fast due to its effective long-range dispersal ability, but from more scattered refugial populations. This, in turn, resulted in a wider distribution of almost all haplotypes, weaker geographic structure and in more evenly distributed haplotype numbers. This hypothesis is in line with the recently observed population expansion of *S. fimbriatum* in several parts of Europe. This species currently colonizes bare surfaces where no other competitive species are present. These sites might be similar to those that remained free after the retreat of the ice sheet and thus *S. fimbriatum* might have recolonized them rapidly. This would imply that good colonizing ability of *S. fimbriatum* is an inherent characteristic of the species and its recent expansion is a response to changing environmental conditions.

References

Aguinagalde I, Hampe A, Mohatny A, Martin JP, Duminil J, Petit RJ. 2005. Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography* **32**: 329-339.

Appelgreen L, Cronberg N. 1999. Genetic and morphological variation in the rare epiphytic moss *Neckera pennata* Hedw. *Journal of Bryology* **21**: 97-107.

Avise JC. 2000. *Phylogeography, the history and formation of species*. Cambridge, MA, USA: Harvard University Press.

Baum DA, Small RL, Wendel JF. 1998. Biogeography and floral evolution of baobabs (Adansonia, Bombacaceae) as inferred from multiple data sets. *Systematic Biology* **47**: 181-207.

Brochmann C, Gabrielsen TM, Nordal I, Landvik JY, Elven R. 2003. Glacial survival or *tabula rasa* ? The history of the north atlantic biota revisited. *Taxon* **52**: 417-450.

Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657-1660.

Crandall KA, Templeton AR. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134**: 959-969.

Cronberg N. 1991. Reproductive biology of *Sphagnum*. *Lindbergia* **17**: 69-82.

Cronberg N. 1996a. Isoenzyme evidence of relationship within *Sphagnum* sect. A*cutifolia* (Sphagnaceae, Bryophyta). *Plant Systematics and Evolution* **203**: 41-67.

Cronberg N. 1996b. Clonal structure and fertility in a sympatric population of the peat mosses *Sphagnum rubellum* and *Sphagnum capillifolium*. *Canadian Journal of Botany* **74**: 1375-1385.

Cronberg N. 1998. Population structure and interspecific differentiation of the peat moss sister species *Sphagnum rubellum* and *S. capillifolium* (Sphagnaceae) in northern Europe. *Plant Systematics and Evolution* **209**: 139-158.

Cronberg N. 2004. Genetic differentiation between populations of the moss *Hylocomium splendens* from low versus high elevation in the Scandinavian mountain range. *Lindbergia* **29**: 64-72.

Daniels RE, Eddy A. 1985. Handbook of European *Sphagna*. Institute of Terrestrial Ecology, UK: Huntingdon.

Daniels RE. 1982. Isozyme variation in British populations of *Sphagnum pulchrum* (Braithw.)Warnst. *Journal of Bryology* **12**: 65-76.

Daniels RE. 1985a. Isozyme variation in populations of *Sphagnum recurvum* var. *mucronatum* from Britain and Finland. *Journal of Bryology* **13**: 563-570.

Daniels RE. 1985b. Isozyme variation in Finnish and British populations of *Sphagnum compactum*. *Annales Botanici Fennici* **22**: 275-279.

Feldmeyer-Christe E, Schnyder N, Bisang I. 2001. Distribution and habitats of peat mosses, *Sphagnum*, in Switzerland. *Lindbergia* **26**: 8-22.

Fernandez CC, Shevock JR, Glazer AN, Thompson JN. 2006. Cryptic species within the cosmopolitan desiccation-tolerant moss *Grimmia laevigata*. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 637-642.

Flatberg KI, Thingsgaard K. 2003. Taxonomy and geography of *Sphagnum tundrae* with a description of *S. mirum*, sp. nov. (Sphagnaceae, sect. *Squarrosa*). *Journal of Bryology* **106**: 501-515.

Forment P, Jovet-Ast S. 1950. *Sphagnum imbricatum* Russ., subfossil en France dans une tourbe des environs de Calais. *Revue Bryologique et Lichenologique* **19**: 125-126.

Frankl R, Schmeidl H. 2000. Vegetation change in a south German raised bog: ecosystem engineering by plant species, vegetation switch or ecosystem level feedback mechanism? *Flora* **195**: 267-276.

Gajewski K, Viau A, Sawada M, Atkinson D, Wilson S. 2001. *Sphagnum* peatland distribution in North America and Eurasia during the past 21,000 years. *Global Biogeochemical Cycles* **15**: 297-310.

Green BH. 1968. Factors influencing the spatial and temporal distribution of *Sphagnum imbricatum* Hornsch. ex Russ. In the British Isles. *Journal of Ecology* **56**: 47-58.

Hamrick JL, Godt MJW. 1996. Effects of life history traits on genetic diversity in plant species. *Phylosophical Transactions of the Royal Society of London Series B* **351**: 1291-1298.

Heuertz M, Fineschi S, Anzidei M, Pastorelli R, Salvini D, Paule L, Frascaria-Lacoste N, Hardy OJ, Vekemans X, Vendramin GG. 2004. Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology* **13**: 3437- 3452.

Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907-913.

Hey J, Wakeley J. 1997. A coalescent estimator of the population recombination rate. *Genetics* **145**: 833-846.

Hudson RR. 1990. Gene genealogies and the coalescent process. In: Harvey PH, Partridge L, eds. *Oxford Surveys in Evolutionary Biology. Vol. 7.* Oxford University Press, USA: New York, 1-44.

Infante M, Heras P. 1987: *Sphagnum imbricatum* Hornsch. Ex Russ., semifosil, en la turbera de Saldropo-Brazar (Ceanuri, Vizcaya). *Estudios Instituto Alaves de la Naturaleza* **2**: 201-203.

Jankowiak K, Rybarczyk A, Wyatt R, Odrzykoski I, Pacak A, Szwejkowska-Kulinska Z. 2005: Organellar inhheritance in the allopolyploid moss *Rhizomnium pseudopunctatum*. *Taxon* **54**: 383-388.

Kooijman AM, Bakker C. 1995. Species replacement in the bryophyte layer in mires – the role of water type, nutrient supply and interspecific interactions. *Journal of Ecology* **83**: 1-8.

Korpelainen H, Laitinen R, Pohjamo M. 2004. Lack of intraspecific variation in cpDNA in *Trichocolea tomentella*. *Jornal of Bryology* **26**: 221-222.

Lascoux M, Palmé AE, Cheddadi R, Latta RG. 2004. Impact of ice ages on the genetic structure of trees and shrubs. *Philosophycal Transactions of the Royal Society of London, Series B* **359**: 197-207.

Longton RE, Schuster RM. 1983. Reproductive biology. In.: Schuster RM, ed. *New Manual of Bryology*. Nichinan, Hattori Botanical Laboratory, 386-461.

Mackay AW, Tallis JT. 1996. Summit-type blanket mire erosion in the forest of Bowland, Lancashire, UK: predisposing factors and implications for conservation. *Biological Conservation* **76**: 31-44.

Malmer N, Wallen B. 1999. The dynamics of peat accumulation on bogs: mass balance of hummocks and hollows and its variation throughout a millenium. *Ecography* **22**: 736-750.

McLachlan *et al*. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* **86**: 2088-2098.

McDaniel SF, Shaw AJ. 2005. Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. *Molecular Ecology* **14**: 1121-1132.

Natcheva R, Cronberg N. 2003. Genetic diversity in the populations of *Sphagnum capillifolium* (Ehrh.) Hedw. from the mountains of Bulgaria, and their possible refugial role. *Journal of Bryology* **25**: 91-99.

Odrzykoski IJ, Sweykowski J. 1991. Genetic differentiation without concordant morphological divergence in the thallose liverwort *Conocephalum conicum*. *Plant Systematics and Evolution* **178**: 135-151.

Pacak A, Szweykowska-Kulinska Z. 2003. Organellar inheritance in liverworts: an example of *Pellia borealis*. *Journal of Molecular Evolution* **56**: 11-17.

Palmé AE, Semerikov V, Lascoux M. 2003a. Absence of geographical structure of chloroplast DNA variation in sallow, *Salix cinerea*. *Heredity* **91**: 465-474.

Palmé AE, Su Q, Rautenberg A, Manni F, Lascoux M. 2003b. Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology* **12**: 201-212.

Petit RJ, Aguinagalde I, de Bealuiei JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohatny A, Müller-Starck G, Demesure-Musch B, Palmé A, Martin JP, Rendell S, Vendramin GG. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**: 1563-1565.

Provan J, Powell W, Hollingsworth PM. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution* **16**: 142-147.

Ramos-Onsins SE, Rosas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**: 2092-2100.

Rozas J, Sanches-Delbarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496-2497.

Såstad SM, Flatberg IK, Hanssen L. 2000. Origin, taxonomy and population structure of the allopolyploid peat moss *Sphagnum majus*. *Plant Systematics and Evolution* **225**: 73-84.

Schröck C, Krisai R. 1999. Verbreitung und Lebensräume ausgewählter *Sphagnum* Arten im Bundesland Salzburg und seinen Nachbargebieten. *Abhandlungen der Zoologisch-Botanischen Gesellschaft in Österreich* **30**: 153-158.

Shaw AJ, Werner O, Ros RM. 2003. Intercontinental mediterranean disjunct mosses: morphological and molecular patterns. *American Journal of Botany* **90**: 540-550.

Shaw AJ, Melosik I, Cox CJ, Boles SB. 2005. Divergent and reticulate evolution in closely related species of *Sphagnum* section Subsecunda. *Bryologist* **108**: 363-376.

Shaw AJ. 2001. Biogeographic patterns and cryptic speciation in bryophytes. *Journal of Biogeogaphy* **28**: 253-261.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **91**: 142-166.

Stenøien HK, Såstad SM. 1999. Genetic structure in three haploid peat mosses (*Sphagnum*). *Heredity* **82**: 391-400.

Sundberg S, Rydin H. 1998. Spore number in *Sphagnum* and its dependence on spore and capsule size. *Journal of Bryology* **20**: 1-16.

Sundberg S, Rydin H. 2002. Habitat requirements for establishment of *Sphagnum* from spores. *Journal of Ecology* **90**: 268-278.

Sundberg S. 2005. Larger capsules enhance short-range spore dispersal in *Sphagnum*, but what happens further away? *Oikos* **108**: 115-124.

Sweykowski J, Buczkowska K, Odrzykoski IJ. 2005. *Conocpehalum salebrosum* (Marchantiopsida, Conocephalaceae) – a new holarctic liverwort species. *Plant Systematics and Evolution* **253**: 133-158.

Swofford DL. 1999. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Michigan.

Szurdoki E, Ódor P. 2005. Distribution and expansion of *Sphagnum fimbriatum* Wils. in Hungary. *Lindbergia* **29**: 136-142.

Taberlet P, Gielly L, Patou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1109.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.

Templeton AR. 2004. Statistical phylogeography: methods of evaulating and minimizing inference errors. *Molecular Ecology* **13**: 789-809.

Templeton AR, Boerwinkle E, Sing CF. 1987. A cladistic analysis of of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics* **117**: 343-351.

Templeton AR, Crandall KA, Sing CF. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. *Genetics* **132**: 619-633.

Templeton AR, Sing CF. 1993. A cladistic analysis of of phenotypic associations with haplotypes inferred from restriction ebndonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* **134**: 659-669.

Thingsgaard K. 2001. Population structure and genetic diversity of the amphiatlantic haploid peat moss *Sphagnum affine* (Sphagnopsida). *Heredity* **87**: 485-496.

Thingsgaard K. 2002. Taxon delimitation and genetic similarities of the *Sphagnum imbricatum* complex, as revealed by enzyme electrophoresis. *Journal of Bryology* **24**: 3-15.

Van der Velde M, Bijlsma R. 2003. Phylogeography of five *Polytrichum* species within Europe. *Biological Journal of the Linnean Society* **78**: 203-213.

Van der Velde M, Van de Zande L, Bijlsma R. 2001. Genetic structure of *Polytrichum formosum* in relation to the breeding system as revealed by microsatellites. *Journal of Evolutionary Biology* **14**: 288-295.

VanAndel TH, Tzedakis PC. 1996. Paleolithic landscapes of Europe and environs, 150,000-25,000 years ago: An overview. *Quaternary Science Reviews* **15**: 481-500.

Wilcox DA, Andrus RE. 1987. The role of *Sphagnum fimbriatum* in secondary succession in a road salt impacted bog. *Canadian Journal of Botany* **65**: 2270-2275.

Acknowledgements

We thank Adam Hölzer, Alain Vanderpoorten, Andrea Fodor, Andrej Martinčic, Anna Ganeva, Baiba Bambe, Beáta Papp, Beatriz Itten, Blanka Shaw, Christian Dolnik, Claes Persson, Erzsébet Szurdoki, Harald Zechmeister, Heike Hofmann, José Pujos, Kai Vellak, ldikó László, Liudmila Abramova, Markus Preussing, Marta Infante, Michael Ignatov, Monika Koperski, Nele Ingerpuu, Patxi Heras, Péter Ódor, Philip De Zuttere, Pierre Goubet, Risto Virtanen, Robert Krisai, Rudolf Soltés, Sam Bosanquet, Sebastian Sundberg, Sigurd Såstad, Sture Westerberg, Zoltán Illyés and Zoltán Tóth for collecting material and/or for field assistance. We are very thankful to Chloé Galley for correcting the language. Valuable comments of three anonymous referees on an earlier draft of the manuscript are also greatly appreciated. This research was financially supported by a Young Academics Grant of the University of Zurich (Forschungskredit der Universität Zürich) and by a Travel Grant of the Swiss National Academy of Sciences.

Table 4 Detailed list of the *Sphagnum fimbriatum* accessions analysed for all three cpDNA regions in this study.

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype	GenBank accession number		
						trnLF	trnSG	rpL16
Sphagnum fimbriatum	15	Norway, Svalbard	150 m	77° 48' N 15° 28' E	A	DQ860276	DQ860278	DQ860277
Sphagnum fimbriatum	120	Spain, Candin, Lamela	850 m	42° 57' N 06° 41' W	B	DQ857522	DQ857554	DQ857418
Sphagnum fimbriatum	368	France, Bretagne	-	48° 13' N 03° 52' W	B	DQ857523	DQ857555	DQ857419
Sphagnum fimbriatum	153	France, Walleis-Auemberg	-	50° 14' N 03° 41' E	C	DQ857524	DQ857551	DQ857415
Sphagnum fimbriatum	261	Spain, Barranco Larreakorta	700 m	43° 00' N 02° 49' W	C	DQ857525	DQ857552	DQ857416
Sphagnum fimbriatum	269	Spain, Candin, Lamela	861 m	42° 57' N 06° 41' W	C	DQ857521	DQ857550	DQ857414
Sphagnum fimbriatum	350	UK, Cors Farlais, Carmarthenshire	300m	51° 54' N 04° 06' W	C	DQ857526	DQ857553	DQ857417
Sphagnum fimbriatum	374	Estonia, Tartu	-	58° 23' N 27° 06' E	D	DQ857528	DQ857596	DQ857460
Sphagnum fimbriatum	2	Scotland, Trinafour	300 m	56° 45' N 03° 55' W	E	DQ857466	DQ857535	DQ857399
Sphagnum fimbriatum	8	Hungary, Also-Erdo I	330 m	47° 24' N 16° 33' E	E	DQ857486	DQ857562	DQ857426
Sphagnum fimbriatum	10	Hungary, Budoskut-Arpadforras	435 m	47° 22' N 16° 28' E	E	DQ857495	DQ857571	DQ857435
Sphagnum fimbriatum	67	Bohemia, Chlum u trebone	-	49° 27' N 13° 35' E	E	DQ857461	DQ857530	DQ857394
Sphagnum fimbriatum	69	Hungary, Velencei-to	-	47° 11' N 18° 33' E	E	DQ857472	DQ857541	DQ857405
Sphagnum fimbriatum	70	Hungary, Szigetcsep Csupics-sziget	-	47° 15' N 18° 59' E	E	DQ857475	DQ857544	DQ857408
Sphagnum fimbriatum	72	Hungary, Dorog	150 m	47° 43' N 18° 45' E	E	DQ857497	DQ857573	DQ857437
Sphagnum fimbriatum	75	Hungary, Monostori-to	345 m	46° 54' N 17° 35' E	E	DQ857503	DQ857579	DQ857443
Sphagnum fimbriatum	79	Austria, Elixhausen	560 m	47° 53' N 13° 01' E	E	DQ857501	DQ857577	DQ857441
Sphagnum fimbriatum	84	Upper Austria, Tarsdorf	480 m	48° 05' N 12° 51' E	E	DQ857463	DQ857532	DQ857396
Sphagnum fimbriatum	93	Germany, Durchenbergried	430 m	47° 45' N 08° 58' E	E	DQ857489	DQ857565	DQ857429
Sphagnum fimbriatum	100	Germany, Hüven	8 m	52° 46' N 07° 34' E	E	DQ857487	DQ857563	DQ857427
Sphagnum fimbriatum	102	Germany, Butterloch	200 m	51° 35' N 10° 18' E	E	DQ857462	DQ857531	DQ857395
Sphagnum fimbriatum	108	Sweden, Mulön	-	65° 37' N 22° 17' E	E	DQ857488	DQ857564	DQ857428
Sphagnum fimbriatum	109	Sweden, Bensbyn	-	65° 38' N 22° 13' E	E	DQ857493	DQ857569	DQ857433
Sphagnum fimbriatum	110	Germany, Neunkirchen	10 m	52° 46' N 08° 44' E	E	DQ857499	DQ857575	DQ857439
Sphagnum fimbriatum	114	Germany, Hamberger Moor	14 m	53° 16' N 08° 51' E	E	DQ857471	DQ857540	DQ857404
Sphagnum fimbriatum	118	France, Troncais forest	280 m	49° 52' N 04° 38' E	E	DQ857490	DQ857566	DQ857430
Sphagnum fimbriatum	132	Belgium, Nonceveux	100 m	50° 26' N 05° 43' E	E	DQ857494	DQ857570	DQ857434
Sphagnum fimbriatum	137	Latvia, sample 4	-	56° 40' N 25° 55' W	E	DQ857469	DQ857538	DQ857402

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype	GenBank accession number		
						trnLF	trnSG	rpL16
Sphagnum fimbriatum	138	Latvia, sample 3	-	56° 40' N 25° 55' W	E	DQ857481	DQ857557	DQ857421
Sphagnum fimbriatum	139	Latvia, sample 2	-	56° 37' N 26° 20' E	E	DQ857498	DQ857574	DQ857438
Sphagnum fimbriatum	145	Germany, Rheinland-Pfalz	120 m	49° 24' N 07° 42' E	E	DQ857467	DQ857536	DQ857400
Sphagnum fimbriatum	150	Hungary, Füzes-to	250 m	47° 02' N 16° 48' E	E	DQ857478	DQ857547	DQ857411
Sphagnum fimbriatum	152	Hungary, Kócse-to	250 m	47° 01' N 16° 49' E	E	DQ857506	DQ857582	DQ857446
Sphagnum fimbriatum	166	Finland, Ruovesi	-	62° 04' N 24° 21' E	E	DQ857464	DQ857533	DQ857397
Sphagnum fimbriatum	169	Finland, Ylikiiminki	-	64° 53' N 26° 07' E	E	DQ857470	DQ857539	DQ857403
Sphagnum fimbriatum	170	Finland, Oulu	-	64° 34' N 25° 33' E	E	DQ857465	DQ857534	DQ857398
Sphagnum fimbriatum	171	Finland, Kuusamo	-	65° 48' N 29° 50' E	E	DQ857484	DQ857560	DQ857424
Sphagnum fimbriatum	172	Finland, Piippola	-	64° 13' N 25° 48' E	E	DQ857500	DQ857576	DQ857440
Sphagnum fimbriatum	173	Belgium, Province of Liege	-	50° 12' N 05° 26' E	E	DQ857477	DQ857546	DQ857410
Sphagnum fimbriatum	174	Belgium, Province of Liege	300 m	50° 14' N 05° 27' E	E	DQ857491	DQ857567	DQ857431
Sphagnum fimbriatum	178	Slovakia, Poprad Basin	680 m	49° 03' N 20° 17' E	E	DQ857505	DQ857581	DQ857445
Sphagnum fimbriatum	236	Sweden, Rödskäret (island)	5 m	59° 36' N 19° 27' W	E	DQ857468	DQ857537	DQ857401
Sphagnum fimbriatum	237	Sweden, Norra Skrakskär (island)	5 m	59° 35' N 19° 20' E	E	DQ857504	DQ857580	DQ857444
Sphagnum fimbriatum	251	Austria, Kneisselmoor	-	47° 45' N 13° 01' E	E	DQ857474	DQ857543	DQ857407
Sphagnum fimbriatum	252	Austria, Zehmemoos	-	47° 59' N 12° 55' E	E	DQ857496	DQ857572	DQ857436
Sphagnum fimbriatum	253	Switzerland, Le Ponten del Mare	1000 m	46° 58' N 06° 43' E	E	DQ857492	DQ857568	DQ857432
Sphagnum fimbriatum	254	Austria, Wengermoor	-	47° 55' N 13° 10' E	E	DQ857502	DQ857578	DQ857442
Sphagnum fimbriatum	257	Hungary, Nyires-to	150 m	47° 49' N 22° 25' E	E	DQ857479	DQ857548	DQ857412
Sphagnum fimbriatum	258	Hungary, Nyires-to	150 m	47° 49' N 22° 25' E	E	DQ857480	DQ857549	DQ857413
Sphagnum fimbriatum	260	France, locality 2005/2	-	45° 51' N 02° 10' E	E	DQ857482	DQ857558	DQ857422
Sphagnum fimbriatum	262	Spain, Pena la Gallina	1622 m	40° 32' N 01° 42' W	E	DQ857473	DQ857542	DQ857406
Sphagnum fimbriatum	266	France, locality 2005/4	-	45° 49' N 01° 59' E	E	DQ857476	DQ857545	DQ857409
Sphagnum fimbriatum	267	Spain, Los Ojos	1307 m	40° 32' N 01° 38' W	E	DQ857483	DQ857559	DQ857423
Sphagnum fimbriatum	268	Spain, Los Ojos	1307 m	40° 32' N 01° 38' W	E	DQ857485	DQ857561	DQ857425
Sphagnum fimbriatum	298	Poland, Silesian lowland	-	51° 06' N 18° 22' E	E	DQ857507	DQ857583	DQ857447
Sphagnum fimbriatum	300	Hungary, Regéc	300 m	48° 26' N 21° 26' E	E	DQ857508	DQ857584	DQ857448
Sphagnum fimbriatum	302	Poland, Beskid Makowski Mont.	425 m	49° 40' N 19° 45' E	E	DQ857509	DQ857585	DQ857449
Sphagnum fimbriatum	304	Hungary, Nagymohos	294 m	48° 15' N 20° 13' E	E	DQ857510	DQ857586	DQ857450
Sphagnum fimbriatum	305	Hungary, Springs of Tegda valley	300 m	48° 23' N 21° 34' E	E	DQ857511	DQ857587	DQ857451
Sphagnum fimbriatum	306	Hungary, Kismohos	296 m	48° 15' N 20° 13' E	E	DQ857512	DQ857588	DQ857452

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype	GenBank accession number		
						trnLF	trnSG	rpL16
Sphagnum fimbriatum	307	Poland, Beskid Slaski Mts.	862 m	49° 38' N 19° 09' E	E	DQ857513	DQ857589	DQ857453
Sphagnum fimbriatum	342	Russia, Rybachy	5 m	55° 09' N 20° 49' E	E	DQ857514	DQ857590	DQ857454
Sphagnum fimbriatum	343	Germany, Jagen	12 m	54° 23' N 09° 32' E	E	DQ857515	DQ857591	DQ857455
Sphagnum fimbriatum	351	UK, Rhos Rydd, Cardiganshire	-	04° 44' N 52° 03' W	E	DQ857516	DQ857592	DQ857456
Sphagnum fimbriatum	357	Sweden, Göteborg, Halland	10 m	57° 31'N 12° 10' E	E	DQ857517	DQ857593	DQ857457
Sphagnum fimbriatum	372	Estonia, Nätsi	-	58° 30' N 24° 04' E	E	DQ857518	DQ857594	DQ857458
Sphagnum fimbriatum	373	Estonia, Tartu	-	58° 23' N 27° 06' E	E	DQ857519	DQ857595	DQ857459
Sphagnum fimbriatum	1	Scotland, Forest of Alyth	400 m	56° 30' N 03° 20' W	F	DQ857527	DQ857556	DQ857420
Sphagnum fimbriatum	375	Estonia, Rapla county	-	59° 02' N 24° 29' E	G	DQ857520	DQ857529	DQ857393

Table 5 Detailed list of the *Sphagnum squarrosum* accessions analysed for all three cpDNA regions in this study.

Species	Number of populations	Locality	Longitude/Latitude	m a.s.l	Haplotype	GenBank accession number		
						trnLF	trnSG	rpL16
Sphagnum squarrosum	265	France, locality 2005/3.	45° 57' N 03° 41' E	1027 m	A	DQ857677	DQ857719	DQ857635
Sphagnum squarrosum	333.1	Montenegro, Durmitor	42° 83' N 18° 23' E	1320 m	A	DQ857678	DQ857720	DQ857636
Sphagnum squarrosum	333.2	Montenegro, Durmitor	42° 83' N 18° 23' E	1320 m	A	DQ857679	DQ857721	DQ857637
Sphagnum squarrosum	354	Russia, Arckhangelsk region, Bolshojslovetski Island	64° 42' N 39° 44' E	-	A	DQ857680	DQ857722	DQ857638
Sphagnum squarrosum	85.1	Austria, Filzmoos	48° 06' N 12° 53' E	480 m	B	DQ857673	DQ857715	DQ857631
Sphagnum squarrosum	85.2	Austria, Filzmoos	48° 06' N 12° 53' E	480 m	B	DQ857675	DQ857717	DQ857633
Sphagnum squarrosum	136	Latvia, locality 8.	56° 37' N 26° 20' E	-	B	DQ857676	DQ857718	DQ857634
Sphagnum squarrosum	158	Finland, Oulu	64° 58' N 25° 56' E	10 m	B	DQ857674	DQ857716	DQ857632
Sphagnum squarrosum	324	Poland, Silesian upland	51° 11' N 18° 38' E	740 m	C	DQ857652	DQ857694	DQ857610
Sphagnum squarrosum	18	Norway, Sogn og Fjordane (THR158499)	62° 00' N 10° 00' E	345 m	D	DQ857640	DQ857682	DQ857598
Sphagnum squarrosum	30	Pyrenees, Lac d`Oredon	42° 49' N 00° 10' E	1900 m	D	DQ857649	DQ857691	DQ857607
Sphagnum squarrosum	96	Germany, Wasenmoos	47° 41' N 09° 35' E	461 m	D	DQ857641	DQ857683	DQ857599
Sphagnum squarrosum	105	Bulgaria, Vitosha mountains	40° 22' N 23° 22' E	1750 m	D	DQ857645	DQ857687	DQ857603
Sphagnum squarrosum	155	Finland, Hyrynsalmi	64° 19' N 29° 00' E	-	D	DQ857648	DQ857690	DQ857606
Sphagnum squarrosum	156	Finland, Vörtsila	61° 56' N 29° 42' E	-	D	DQ857643	DQ857685	DQ857601
Sphagnum squarrosum	160	Finland, Vieremä	63° 19' N 26° 32' E	-	D	DQ857644	DQ857686	DQ857602
Sphagnum squarrosum	175	Slovakia, High Tatra mountains	20° 03' N 49° 07' E	1335 m	D	DQ857642	DQ857684	DQ857600
Sphagnum squarrosum	176	Slovakia, High Tatra mountains	49° 12' N 20° 16' E	1079 m	D	DQ857646	DQ857688	DQ857604
Sphagnum squarrosum	289	Austria, Seemoos	47° 10' N 13° 47' E	1673 m	D	DQ857650	DQ857692	DQ857608
Sphagnum squarrosum	292	Romania, Kelemen havasok	47° 14' N 25° 16' E	800 m	D	DQ857651	DQ857693	DQ857609
Sphagnum squarrosum	335	Serbia, Golija	43° 20' N 20° 15' E	1484 m	D	DQ857639	DQ857681	DQ857597
Sphagnum squarrosum	81.1	Austria, Lungau	48° 06' N 12° 52' E	1260 m	D	DQ857647	DQ857689	DQ857605
Sphagnum squarrosum	134	Russia, Moscow	55° 30' N 37° 30' E	-	E	DQ857654	DQ857696	DQ857612
Sphagnum squarrosum	147	Hungary, Velencei-tó	47° 12' N 18° 33' E	100 m	E	DQ857655	DQ857697	DQ857613
Sphagnum squarrosum	239	Sweden, Långmossen	59˚ 58’ N 17˚ 18’ E	0 m	E	DQ857653	DQ857695	DQ857611
Sphagnum squarrosum	360	Göteborg, Rammsjödal	57° 31' N 12° 10' E	10 m	E	DQ857656	DQ857698	DQ857614
Sphagnum squarrosum	101	Germany, Hüven	52° 46' N 07° 34' E	8 m	F	DQ857672	DQ857714	DQ857630

Species	Number of populations	Locality	Longitude/Latitude	m a.s.l	Haplotype	GenBank accession number		
						trnLF	trnSG	rpL16
Sphagnum squarrosum	141	Latvia, locality 5.	56° 37' N 26° 20' E	-	F	DQ857671	DQ857713	DQ857629
Sphagnum squarrosum	161	Finland, Hailuoto	65° 05' N 24° 46' E	10 m	F	DQ857670	DQ857712	DQ857628
Sphagnum squarrosum	103	Germany, Butterloch	51° 35' N 10° 18' E	200 m	G	DQ857658	DQ857700	DQ857616
Sphagnum squarrosum	107	Sweden, Mulön	65° 37' N 22° 17' E	-	G	DQ857659	DQ857701	DQ857617
Sphagnum squarrosum	112	Germany, Göttingen	51° 34' N 10° 08' E	5 m	G	DQ857661	DQ857703	DQ857619
Sphagnum squarrosum	135	Latvia, locality 7.	56° 40' N 25° 55' E	-	G	DQ857660	DQ857702	DQ857618
Sphagnum squarrosum	140	Latvia, locality 1.	56° 40' N 25° 55' E	-	G	DQ857662	DQ857704	DQ857620
Sphagnum squarrosum	143	Estonia, Nigula	24° 40' N 58° 01' E	10 m	G	DQ857663	DQ857705	DQ857621
Sphagnum squarrosum	159	Finland, Piipola	64° 22' N 25° 40' E	25 m	G	DQ857657	DQ857699	DQ857615
Sphagnum squarrosum	241	Sweden, Rödskäret (island)	59˚ 36’ N 19˚ 27’ E	2 m	G	DQ857664	DQ857706	DQ857622
Sphagnum squarrosum	264	Spain, locality 2005/9.	43° 02' N 02° 42' W	642 m	G	DQ857665	DQ857707	DQ857623
Sphagnum squarrosum	323	Hungary, Tegda	48° 25' N 21° 29' E	440 m	G	DQ857666	DQ857708	DQ857624
Sphagnum squarrosum	334	Serbia, Golija 03/105/6	43° 20' N 20° 15' E	1484 m	G	DQ857667	DQ857709	DQ857625
Sphagnum squarrosum	344	Germany, Jagen	54° 23' N 09° 31' E	12 m	G	DQ857668	DQ857710	DQ857626
Sphagnum squarrosum	356	Russia, Kola Peninsula	68° 05' N 39° 50' E	-	H	DQ857669	DQ857711	DQ857627

Chapter II.

New primers for amplifying the GapC gene of bryophytes and its utility in infraspecific phylogenies in the genus *Sphagnum*

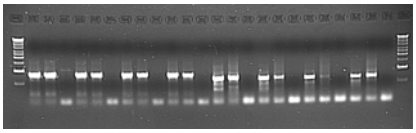
Péter Szövényi[1,2], Zsófia Hock[1,2], Edwin Urmi[1]and Jakob Schneller[1]

[1] *Institute of Systematic Botany, University of Zürich, Zollikerstr. 107, CH-8008 Zürich, Switzerland*

[2] *Cryptogam Laboratory, Department of Plant Taxonomy and Ecology, Eötvös Loránd University, Faculty of Science, Pázmány P. sétány 1/C, H-1117 Budapest, Hungary*

Published in:

Lindbergia 31: 32-39. 2006.



Abstract

Bryophytes show several special features, such as haploid gametophytes and the ability to disperse over long distances, which suggest a different spatial partition of neutral genetic variation than in vascular plants. There are, however, only few nuclear markers for reconstructing infraspecific phylogenies in bryophytes. We designed primers spanning the 2-9 exons and intervening introns of the GapC gene (2550 bp). Our primers were designed to selectively amplify sequences of the genus *Sphagnum*, however, their degeneracy makes it possible to obtain sequences from other groups of bryophytes. In a pilot study, well-resolved infraspecific phylogenies have been obtained using 10 European *Sphagnum fimbriatum* Wils. accessions.

Introduction

Infraspecific phylogeographic studies usually employ uniparentally inherited markers, such as chloroplast and mitochondrial markers. Chloroplast microsatellites are useful markers in analyses at a finer spatial scale, but their homology is hard to assess (Provan et al. 2001). In the recent years, an increasing number of low or single copy nuclear genes have been used in nuclear infraspecific phylogenies in plants (Caicedo & Schaal 2004).

Bryophytes form a diverse group of early land plants (Shaw & Renzaglia 2004). They have traditionally been regarded as evolutionary stenotypic because they are haploid, frequently reproduce vegetatively, and have large distribution areas that often span several continents. Recent studies on the population genetics of bryophytes using isozymes, DNA fingerprinting and to a limited extent nuclear genes have challenged this view and shown that bryophytes are evolutionally much more flexible than earlier believed (Shaw 2001).

Studies on vascular plants using molecular markers have increased during the last decade and have led to different ideas about the effect of ice ages or other historical events on the spatial distribution of genetic diversity in Europe (Petit et al. 2005). Among bryophytes, peat mosses are the most intensively studied group using molecular markers. However, Europe-wide phylogeographic investigations are few and have only been conducted using isoenzyme or fingerprinting methods (Cronberg 1998; Stenøien and Såstad 1999; Thingsgaard 2001).

Peat mosses, like other bryophytes, are capable of long-range dispersal due to several special features, such as numerous spores, and tolerance of both desiccation and freezing. This suggests that this group could have a different pattern of neutral genetic variation than is found in flowering plants. In addition, peat mosses are able to develop a long-term persistent spore bank and to survive over unfavourable periods that can be decades or even centuries in duration (Sundberg & Rydin 2000).

Hitherto, little has been known about the possible glacial refugia of European bryophyte species. The role of the Balkan Mountains as refugia has been suggested by recent investigations (Natcheva & Cronberg 2003; Van der Velde & Bijlsma 2003). However, species with a recent circumboreal distribution are likely to have survived the glacial cycles further north (Cronberg 2004).

This paper represents the preliminary results of an ongoing project investigating the phylogeographic pattern of four *Sphagnum* species using different molecular methods, including AFLPs and sequencing of nuclear and chloroplast markers. In the present paper we

explore the utility of a single copy nuclear gene (GapC), which was selected to study the infraspecific phylogeographic structure of the peat mosses (*Sphagnum*) at a European scale. Genes of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene family encode the homologous NAD^+ dependent GapC (in the cytosol), GapCp (in non green plastids) and the $NADP^+$ dependent GapAB (in chloroplasts) enzymes. Sequences of the GapA, B and C genes are divergent, while the recently discovered GapCp gene shows a high sequence similarity with the GapC gene (Meyer-Gauen et al. 1994, Petersen et al. 2003). GapC genes may occur in several copies, but in bryophytes only one copy has been discovered to date (Petersen et al. 2003). The GapC gene has been used as a molecular marker in several groups of organisms but was only rarely applied to reconstruct infraspecific phylogenies in plants (Figge et al. 1999). One part of the GapC gene (gpd after Wall 2002) has been used in phylogenetic reconstruction of a recently diverged group of bryophytes and yielded well-resolved trees. Wall (2002) provided primers spanning the 5-9 exons, however, in our preliminary studies his primers preferentially amplified fungal contaminants of *Sphagnum* species. Therefore we attempted to design new bryophyte specific primers for amplifying the whole GapC gene and test the gene`s utility in resolving infraspecific phylogenies.

Methods

Primer design

We downloaded and manually aligned sequences of the GapC gene from the GenBank using a text editor. Sequences of the contaminant fungus and the recently discovered GapCp were also included to prevent amplification of contaminants. The downloaded representatives included *Physcomitrella patens* (X72381), *Arabidopsis thaliana* (M64119), *Marchantia polymorpha* (AJ246023), *Sphagnum cuspidatum* (AJ246021), *Glomerella cingulata* (M93427), *Neurospora crassa* (U56397), *Cochliobolus lunatus* (X58718), *Sphagnum cuspidatum* GapCp (AJ246022) sequences. Based on the aligned sequences, first degenerate primers annealing to the 2nd and 9th exon positions were designed. In this first step, primers were constructed to amplify both moss and liverwort sequences but to exclude fungal contaminants and the GapCp gene. Products of accessions 70., 72., 114. and 79. were sequenced using the primers GapC exon2 forw. and exon9 rev. and internal sequencing primers were designed (Table 1 and Fig. 1). Because sequencing of these three individuals of *S. fimbriatum* did not show variability in the region spanning exons 2-5, but showed a considerable amount of variability in the region between the 5th-9th exons only the latter region was used in the following analyses.

DNA extraction, PCR conditions and sequencing

The total genomic DNA was extracted from air and silica gel dried specimens using the Qiagen Dneasy Plant Mini Kit (Qiagen) following the manufacturer`s protocol. PCR amplifications were performed using the primer pairs GapC exon4 forw. and exon9 rev. In case of herbarium specimens usually a second PCR was done using 1 µl from the 10x diluted first round product as template with primers exon5 forw. and exon9 rev. The PCR mix contained 1-2 µl of template DNA (3-5 ng/µl), 11.5 µl ddH_2O, 0.6 µl of each primer (10 µM), 3.2 µl of a 1.25 mM dNTP solution in an equimolar ratio, 0.1 µl Taq DNA-polymerase (5u/µl, Sigma), 2 µl of the provided 10x enzyme buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM $MgCl_2$, 0.01 % gelatin). The PCR cycling conditions were 94 ºC denaturating temperature for 3 min, different annealing temperatures (GapC exon9 rev. and exon4 forw.: 61 ºC; GapC exon9 rev. and exon5 forw.: 55 ºC) for 1 min and 72 ºC extension for 2 min for 35 cycles with a final extension at 72 ºC for 7 min. The products were cleaned using the GFX PCR and gel

band purification kit (Amersham Biosciences), and approximately 10 ng of product was sequenced in 10 µl cycle sequencing reaction with the Big dye v3.1 in an ABI prism 3100 (Applied Biosystem) genetic analyzer.

Phylogenetic analysis

To test the utility of sequences in reconstructing infraspecific phylogenies, 10 specimens of *S. fimbriatum* were selected from different parts of Europe and used in a pilot study (Table 2). Sequences were aligned by eye directly in NEXUS format. The gap coded aligned sequences were subjected to maximum parsimony analysis using the exhaustive search option of PAUP* (Swofford 1999). Bootstrap analysis was performed with 1000 bootstrap replicates, with random addition of sequences (10 replicates, number of trees held at each step = 100). TBR and MulTrees options were used during branch swapping.

Table 1 Names, sequences and usage of GapC primers designed.

Primer name	Usage	Sequence (5´-3´)	Design based on
GapCexon2 forw.	GapC primer	GACGYATYGGCCGCTTGGTRGC	GenBank sequences
GapCexon4 forw.	GapC primer	GGTGGYAAGCCTGTWGCTGTTTAT	GenBank sequences
GapCexon5 rev.	GapC primer	GTTTKCCCCAYGGTATCTCWGAA	GenBank sequences
GapCexon9 rev.	GapC primer	CTATCAGTRATGAAGTCRGTGGAC	GenBank sequences
GapCexon5 forw.	GapC internal and sequencing primer	TGGTGTATTCACAGACAAAGCAA	*Sphagnum* sequences
GapCexon7 forw.	GapC internal sequencing primer	AATGACAAATTTGGAATTGTTG	*Sphagnum* sequences
GapCexon8 rev.	GapC internal sequencing primer	ATGTTGACTCCAGCAGCACGG	*Sphagnum* sequences

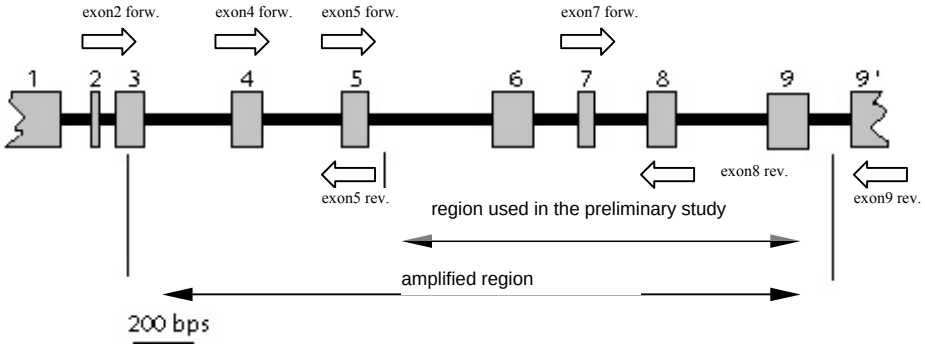


Fig. 1 Structure of the GapC gene in the *Sphagnum* species investigated. Gray boxes represent the 10 exons (see numbers at the bottom of the figure), while solid dark lines the introns between them. Arrows indicate the positions of primers designed. Positions of forward primers are indicated above, reverse primers below the gene. Note that because of a putative intronic sequence the primer exon9 rev. anneals at exon 9‘.

Results

Structure of the GapC gene in the genus Sphagnum

Blast searches (Altschul et al. 1997) verified that only GapC sequences of *Sphagnum* were amplified. Amplification of contaminants was never observed. Sequencing chromatograms showed no ambiguous bases or double peaks, which may indicate that only one copy of the gene was present. Intron-exon boundaries were determined comparing sequences with the GapC sequence of *Arabidopsis thaliana* and the cDNA of *Physcomitrella patens*. The partial GapC gene sequence spanning the 2-9 exons was successfully amplified from *S. fimbriatum, girgensohnii, squarrosum, teres, palustre, centrale, magellanicum* and

Table 2 Codes, locality and GenBank accession numbers of each specimen used in this study.

Code	Species	Name	Region	Herbarium code	GenBank accession
79.	*S. fimbriatum*	Austria	Salzburg county, Ursprung-Moor bei Elixhausen	-	AM088014
132.	*S. fimbriatum*	Belgium	Province of Liège, Nonceveux	-	AM088015
114.	*S. fimbriatum*	Germany	Lower Saxony, Hamberger Moor	-	AM088018
69.	*S. fimbriatum*	Hungary1	Fejér county, Velencei-to	-	AM088013
70.	*S. fimbriatum*	Hungary2	Pest county, Szigetcsép, Csupics-sziget	-	AM088016
72.	*S. fimbriatum*	Hungary3	Pest county, Dorog	-	AM088017
15.	*S. fimbriatum*	Norway	Svalbard, Nordenskiöld land	THR 158202	AM088019
120.	*S. fimbriatum*	Spain1	Galicia, Candin, Lamela	-	AM088010
121.	*S. fimbriatum*	Spain2	Guipuzcoa, Onate, Sierra de Elguea-Urkilla	-	AM088011
261.	*S. fimbriatum*	Spain3	Basque country, Barranco Larreakorta	-	AM088012

papillosum but the complete sequence was only determined for four accessions of *S. fimbriatum* (70., 72., 79., 114. see in Table 2). The amplified region was 2551 bps long in the *S fimbriatum* accessions investigated. The sequence contains a part of the 3rd exon, the entire 4-9 exons with the intervening introns and a partial sequence of the 9th intron (Fig. 1). Although the GapC gene usually contains 9 exons (Wall 2002) in our species the 9th exon is disrupted into two parts (referred to as exon 9 and 9') by an approximately 400 bps long sequence assumed to be noncoding (referred to as intron 9).

Polymorphism of the investigated region

As mentioned previously, only the region spanning the 5th-9th exons was used for the phylogeographic analysis (Fig. 1). This region was 1470-1477 bps long after trimming 50 bases from both ends of the contiged sequences to avoid counting ambiguous bases. The sequences analysed contain the entire region spanning exon 6 - exon 9 and partial sequences of intron 5 and intron 9 (Fig. 1). Exon 7 turned out to be the shortest one (56 bps) while exon 6 was the longest (149 bps). Lengths of the remaining exons were between these two values (exon 8 = 98 bps and exon 9 = 143 bps). Using the ten individuals of *S. fimbriatum* originating from different parts of Europe, all sequences of all accessions turned out to be different, except isolate 70. and 72., which were identical (Table 3). The total length of noncoding sequences ranged from 1024 to 1031 bps, while coding sequences spanned 446 bps. Including all accessions and all types of mutations (both mismatches and indels), 60

Table 3 Summary of polymorphic sites in the region (exon 5 - exon 9) studied. Individual accessions are represented by their codes in the first column (see collecting data of each accession in table 2.). The number of polymorphic sites and the exon-intron structure of the gene are indicated at the top of the figure. The consensus (cons.) of all sequences is given above. Deletions are represented by asterisks (*). Hyphen (-) indicates no change.

intr. 5	ex. 6	intr. 6	ex. 7	intr. 7	ex. 8	intr. 8	ex. 9	intr. 9

	25	28	42	73	100	137	274	358	416	422	428	479	567	591	613	621	632	634	661	709	716	720	726	730	744	757	758	769	770	771
Cons.	G	G	T	T	C	A	T	T	C	T	C	T	C	*	C	A	G	C	T	C	C	A	C	A	G	C	C	A	*	A
120.	-	-	-	-	-	-	-	-	-	C	-	C	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	C	-
121.	-	-	-	-	-	-	-	-	-	C	-	C	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	C	-
261.	-	-	-	-	-	-	-	-	-	C	-	C	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	-	C	-
79.	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	*	-	-
69.	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	-	-
132.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	*
70.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	*
72.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	*
114.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-	-	-	-	-	*
15.	A	C	C	C	G	*	C	C	T	C	-	C	G	-	T	-	-	T	C	T	G	G	T	-	A	A	*	-	C	*

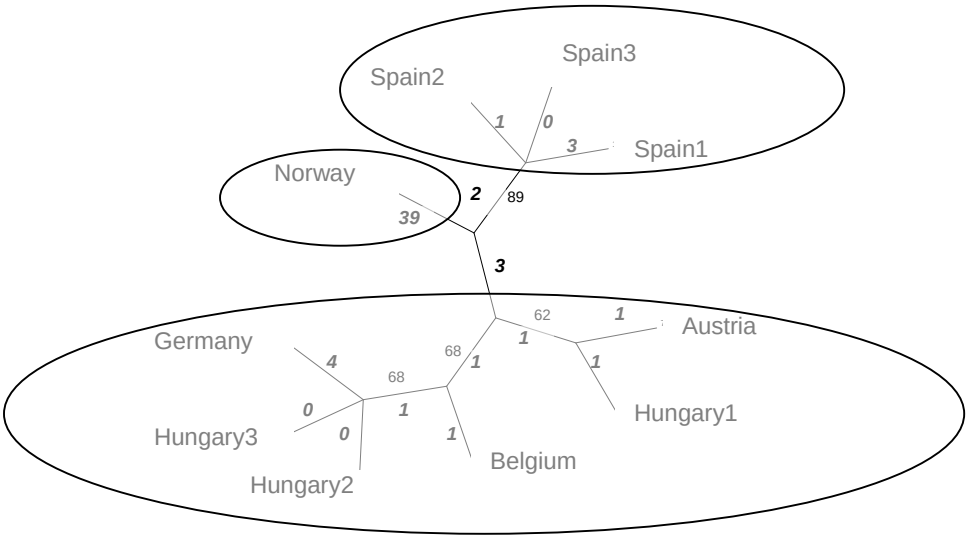
	772	818	819	850	912	914	926	929	932	963	982	1008	1018	1030	1033	1045	1080	1096	1120	1140	1154	1157	1178	1194	1195	1269	1296	1383	1453
Cons.	G	*	T	G	A	G	A	C	T	A	G	T	G	T	*	G	T	T	T	G	G	T	C	G	T	G	G	*	A
120.	-	-	-	-	C	A	-	-	-	-	-	-	-	-	T	-	-	C	-	-	-	-	-	-	-	-	T	-	-
121.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	C	-	-	-	-	-	-	-	-	T	-	-
261.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	T	-	-
79.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
69.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
70.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
72.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
114.	-	-	-	-	-	-	-	-	A	T	-	-	-	A	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15.	*	C	*	C	-	-	G	G	-	-	T	A	A	-	-	A	-	C	A	T	A	*	A	*	*	A	-	T	G

variable sites were found from which 28 occurred in noncoding and 17 in coding regions from which 9 turned out to be parsimony informative. No indels were found in the coding sequences and substitutions caused 6 amino acid changes. All exons except exon 9 contained about the same amount of mutations, which bear relatively a low number of polymorphic sites compared to its length. From the 28 polymorphic sites in introns 15 were indels and 13 substitutions. Single base indels were frequently found in the introns and a two base pair long indel was observed only once. Most of the variation was found in intron 7 and 8. The number of variable sites changed considerably when the strongly differentiated accession from Svalbard was excluded (only 19 variable sites remained from which 9 were parsimony informative).

Our analysis yielded nine equally parsimonious trees. The strict consensus of the nine trees revealed three well-supported (bootstrap value: 89 %, CI=0.9000, RI=0.9500) and easily interpretable clades with Central, North and South European distributions (Fig 2.).

Our analysis revealed three relatively well differentiated infraspecific European clades in *S. fimbriatum*. The accession from Svalbard shows strong sequence divergence and is separated from the other European specimens. The second clade contains accessions from Spain with inter-clade differentiation. The rest of the accessions can be joined in one clade containing accessions originating mostly from central and central eastern Europe. The three major clades are well-supported (bootstrap value 89), whereas the subclades within the Central and Central Eastern European clades are only weakly supported (bootstrap values < 70).

Fig. 2 Strict consensus of 9 equally most parsimonious unrooted cladograms. Numbers in bold along branches show number of nucleotide substitutions. Numbers in plain text indicate bootstrap supports after 1000 replicates. Circles delimit the three major clades. Localities for each accession are given in table 2.



Discussion

Structure and variability of the GapC gene

This study examined the utility of a portion of the GapC gene of bryophytes at the infraspecific level. Nuclear markers appropriate for resolving relationships among closely related species and reconstructing infraspecific phylogenies in bryophytes are scarce. The utility of only a limited number of nuclear markers has been investigated so far, mainly in species level phylogenies (Wall 2002, 2005, Shaw et al. 2003). At the infraspecific level, variability of accessions with only a limited geographic origin has been tested. Several regions investigated spanned only portions of a gene or their evolution was complicated by the presence of several paralogs (Vanderpoorten et al. 2003, McDaniel and Shaw 2005). Heretofore, in bryophytes, utility of the GapC gene in a phylogenetic context has only been investigated in the genus *Mitthyridium* (Wall 2002, 2005). Infraspecific variability of this gene or its relatives has been used in several flowering plants (Olsen and Schaal 1999, Olsen 2002, Isoda and Shiraishi 2001, Tani et al. 2003, Pérusse and Schoen 2004). Furthermore, this gene has been used to resolve relationship and to date the relative age of main green plant lineages (Martin et al. 1993).

In contrast to the gene recovered by Wall (2002) in *Mitthyridium* species, the GapC gene of the genus *Sphagnum* shows structural differences. The 9th exon is disrupted by an approximately 450 bp long sequence assumed to be noncoding. Similar loss of introns in the GapC gene was found in the family Boraginaceae (Pérusse and Schoen 2004), though the structure of the gene seems to be conserved across a wide range of organisms (Martin et al. 1993). Except for the additional noncoding sequence, which causes a considerable increase in the length of the gene, the structure of the GapC gene is very similar to that of *Arabidopsis* (Shih et al. 1991) and *Mitthyridium* (Wall 2002). Proportions of noncoding sequences were higher in *Sphagnum* than in the previously studied bryophyte group, however, this difference was mainly caused by the additional intron disrupting exon 9. Recent investigations confirmed the occurrence of GapC paralogs in several plant families (Pérusse and Schoen 2004). Based on our own investigations and other research have shown (Wall 2002, 2005, Petersen et al. 2003) that the bryophytes that have been investigated so date, only contain a single copy of the gene. It is evident that unambiguous electropherograms cannot by themselves be used to determine the single or multi copy state of the gene, thus confirmation of the copy number by cloning is in progress. Preliminary results showed no difference among

sequences obtained from five clones of one accession, which confirms the single copy hypothesis.

This study provides the first insights into the infraspecific variability of the GapC gene in bryophytes. Infraspecific variability of the gene has been used to study population structure and hybridisation in some flowering plants. All of these studies, even those using only smaller parts of the gene, reported relatively high level of variability among and within populations (Olsen and Schaal 1999, Olsen 2002, Isoda and Shiraishi 2001, Tani et al. 2003, Pérusse and Schoen 2004). In the ten *S. fimbriatum* accessions investigated, introns were a rich source of variability, which is in agreement with the results of previous studies. More than half of the polymorphisms found in introns were indels. The number of indels was also high in infraspecific studies on *Manihot esculenta* (Olsen and Schaal 1999, Olsen 2002) and in phylogenetic studies in the genus *Mitthyridium* (Wall 2002, 2005). In our data set, with one exception, only single base deletions or insertions were detected. As compared to other studies, these insertions and deletions were relatively short. This difference might be caused by the limited number of accessions investigated or it might reflect different evolutionary patterns in the genus *Sphagnum* than in other green plants. Comparing our results to other studies, a higher percent of polymorphic sites were found in coding regions, but only about one third of these caused amino acid changes. In a detailed study on *Manihot esculenta* (Olsen and Schaal 1999) only one replacement substitution was found. This contrasting pattern of exon evolution might reflect cryptic speciation in *S. fimbriatum*, although almost all amino acid replacements occurred with amino acids having the same characteristics.

Interpretation of the preliminary phylogeographic pattern

Our pilot study indicates that the second part of the GapC gene is a useful tool in resolving infraspecific relationships. Amplification of the region is straightforward and sequences can even be obtained from 25 years old herbarium specimens when using the primers in a nested design. We are aware that ten accessions are not enough to make robust phylogeographic conclusions, but some hypotheses to be tested using more detailed sampling can still be drawn. Strong sequence divergence of the accession from Svalbard is evident and implicates a long-standing gene flow barrier, which separated it from the rest of the European lineages. Based on our results, it is not possible to decide the place of origin of the Svalbard lineage. However, it seems very likely, that it originated from another pool than the European. The divergence of the Svalbard accession is supported not only by the sequence infromation,

but also by the macroscopic habit of the plants, which is different from that of the European plants.

We found relatively high amount of sequence divergence between closely related populations in Spain, supporting a hypothesis of glacial survival of the species in the Iberian Peninsula. Survival in isolated, probably small populations increased the effect of genetic drift, which led to fairly strong differentiation between populations.

The third, relatively heterogeneous clade contains central and eastern European lineages. The three Hungarian accessions included seem to be divergent: one of them is more closely related to the accession from Austria than to the other Hungarian ones. These accessions are identical and very similar to the lineage from Belgium. The relatively high divergence among Hungarian lineages is hard to interpret based on our preliminary data set. It might represent survival of the species in different refugia in periglacial areas like the Carpathian Basin. Sheltered tree populations could have served as refugia during the last glacial maximum (Willis, Rudner and Sümegi 2000; Magyari et al. 2000). Another explanation might be that populations of the species survived in different refugia outside the Carpathian Basin and the relatively high divergence among Hungarian lineages is due to an intermingling of formerly isolated populations. Knowing that *S. fimbriatum* is probably spreading in Hungary, Austria and Germany (Szurdoki and Ódor 2004) this second explanation seems to be more likely. The accession from northern Germany has also to be mentioned. This lineage is genetically well differentiated from all other accessions and probably represents a frequent haplotype of northern Europe, however our present data are not detailed enough to make reliable conclusions.

Our preliminary study has shown that the second part of the GapC gene is a promising marker in resolving infraspecific phylogenies. In addition, it might be useful in resolving relationships among closely related species or species complexes. Analysing more accessions would give the opportunity to study the evolution of the GapC gene in more detail and clarify whether cryptic molecular evolution occured in *S. fimbriatum*.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. – Nucleic Acids Res. 25:3389-3402.

Caicedo, A.L. and Schaal, B.S. 2004. Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. – Mol. Ecol. 13: 1871-1882.

Cronberg, N. 1998. Population structure and interspecific differentiation of the peat moss sister species *Sphagnum rubellum* and *S. capillifolium* (Sphagnaceae) in northern Europe. – Plant Syst. Evol. 209: 139-158.

Cronberg, N. 2004. Genetic differentiation between populations of the moss *Hylocomium splendens* from low versus high elevation in the Scandinavian mountain range. – Lindbergia 29: 64-72.

Figge, R.M., Schubert, M., Brinkmann, H. and Cerff, R. 1999. Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eucaryotes: Evidence for intra- and interkingdom gene transfer. – Mol. Biol. Evol. 16: 429-440.

Isoda, K. and Shiraishi, S. 2001. Allelic sequence polymorphisms in the intron region of the nuclear-encoded GapC gene preceded the speciation of three closely related *Abies* species (Pinaceae). – Theor. Appl. Genet. 102: 244-250.

Magyari, E., Jakab, G., Sümegi, P., Rudner, E. and Molnár, M. 2000. Paleobotanical investigations on Mohos-lakes, Kelemér, N-Hungary. In: Szurdoki, E. (ed.), Peat moss habitats in Hungary: Research, Management and Conservation. CEEWEB Munkacsoport, Miskolc, Budapest, pp. 101-133.

Martin, W., Lydiate D., Brinkmann, H., Forkmann, G., Saedler, H. and Cerff, R. 1993. Molecular phylogenies in angiosperm evolution. – Mol. Biol. Evol. 10(1): 140-162.

McDaniel, S.F. and Shaw, A.J. 2005. Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. – Mol. Ecol. 14: 1121-1132.

Meyer-Gauen, G., Hernrand, H., Pahnke, J., Cerff, R. and Martin, W. 1994. Gene structure, expression in *Escherichia coli* and biochemical properties of the NAD^+-dependent glyceraldehyde-3-phosphate dehydrogenase from *Pinus sylvestris* chloroplasts. – Gene 209: 167-174.

Natcheva, R. and Cronberg, N. 2003. Genetic diversity in the populations of *Sphagnum capillifolium* (Ehrh.) Hedw. from the mountains of Bulgaria, and their possible refugial role. – J. Bryol. 25: 91-99.

Olsen, K.M. 2002. Population history of *Manihot esculenta* (Euphorbiaceae) inferred from nuclear DNA sequences. – Mol. Ecol. 11: 901-911.

Olsen, K.M. and Schaal, B.A. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. – Proc. Natl. Acad. Sci. USA 96: 5586-5591.

Pérusse, J.R. and Schoen, D.J. 2004. Molecular evolution of the *GapC* gene family in *Amsinckia spectabilis* populations that differ in outcrossing rate. – J. Mol. Evol. 59: 427-436.

Petersen, J., Brinkmann, H., and Cerff, R. 2003. Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. – J. Mol. Evol. 57: 16-26.

Petit, R.J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D. and Vendramin, G.G. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. – Mol. Ecol 14: 689-701.

Provan, J., Powell, W. and Hollingsworth, P.M. 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. – Trends. Ecol. Evol. 16(3): 142-147.

Shaw, A.J. 2001. Biographic patterns and cryptic speciation in bryophytes. – J. Biogeogr. 28: 253-261.

Shaw, A.J., Cox, C.J. and Boles, S.B. 2003. Polarity of peatmoss (*Sphagnum*) evolution: who says bryophytes have no roots? – Am. J. Bot. 90(12): 1777-1787.

Shaw, J., Renzaglia, K. 2004. Phylogeny and diversification of bryophytes. – Am. J. Bot. 91: 1557-1581.

Shih, M.C., Heinrich, P. and Goodman, H.M. 1991. Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate-dehydrogenase from *Arabidopsis thaliana*. – Gene 104: 133-138.

Stenøien, H. K. and Såstad, S. M. 1999. Genetic structure in three haploid peat mosses (*Sphagnum*). – Heredity 82: 391-400.

Sundberg, S. and Rydin, H. 2000. Experimental evidence for a persistent spore bank in *Sphagnum*. – New Phytol. 148: 105-116.

Swofford, D.L. 1999. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Suderland, Michigan.

Szurdoki, E. and Ódor, P. 2004. Distribution and expansion of *Sphagnum fimbriatum* in Hungary. – Lindbergia 29: 136-142.

Tani, N., Tsumura, Y. and Sato, H. 2003. Nuclear gene sequences and DNA variation of *Cryptomeria Japonica* samples from the postglacial period. – Mol. Ecol. 12: 859-868.

Thingsgaard, K. 2001. Population structure and genetic diversity of the amphiatlantic haploid peat moss *Sphagnum affine* (Sphagnopsida). – Heredity 87: 485-496.

Van der Velde, M. and Bijlsma, R. 2003. Phylogeography of five *Polytrichum* species within Europe. – Biol. J. Linn. Soc. 78: 203-213

VanAndel, T.H. and Tzedakis, P.C. 1996. Paleolithic landscapes of Europe and environments, 150,000-25,000 years ago: An overview. – Quaternary Science Reviews 15: 481-500.

Vanderpoorten, A., Shaw, A.J. and Cox C.J. 2003. Evolution of multiple paralogous adenosine kinase genes in the moss genus *Hygroamblystegium*: phylogenetic implications. – Mol. Phylogenet. Evol. 31: 505-516.

Wall, D.P. 2002. Use of the nuclear gene glyceraldehydes 3-phosphate dehydrogenase for phylogeny reconstruction of recently diverged lineages in *Mitthyridium* (Musci: *Calymperaceae*). – Mol. Phylogenet. Evol. 25: 10-26.

Wall, D.P. 2005. Origin and rapid diversification of a tropical moss. – Evolution 59(7): 1413-1424.

Willis, K.J., Rudner, E. and Sümegi, P. 2000. The full-glacial forests of central and southeastern Europe: Evidence from Hungarian palaeoecological records. – Quaternary Research 53: 203-213.

Acknowledgements

We are very thankful to the following people who kindly provided us plant material or accompanied us during field work (names are given in an alphabetic order): Alain Vanderpoorten, Andrea Fodor, Marta Infante, Monika Koperski, Patxi Heras, Robert Krisai and, Sigurd Såstad.We are also grateful to two anonymous reviewers for critical edits. This research is finacially supported by the “Young Academics Support of the University of Zürich” (Forschungskredit der Universität Zürich).

Chapter III.

Contrasting patterns of multilocus genetic structure of two peat moss species in Europe: influence of demography, mating system and historical factors

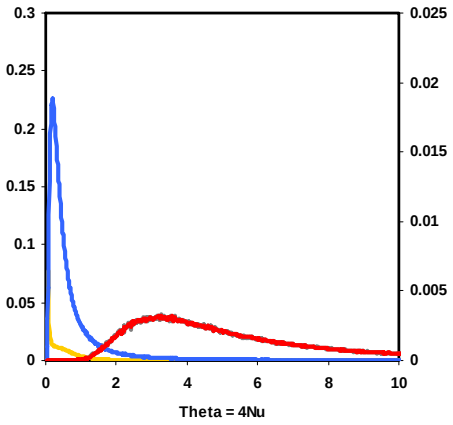
Péter Szövényi[1,2], Zsófia Hock[1,2], Johann Jakob Schneller[1], Zoltán Tóth[2]

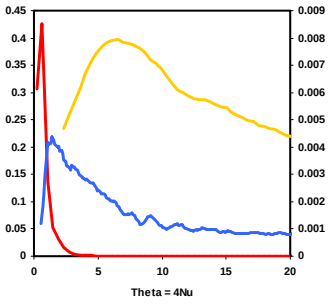
[1]*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, Zurich 8008, Switzerland*

[2]*Eötvös Loránd University, Department of Plant Taxonomy and Ecology, Pázmány Péter sétány 1/C, Budapest 1117, Hungary*

In review:

Molecular Ecology





Abstract

We investigated the amount and pattern of sequence variability of two peat moss species with similar distributions and mating systems but presumably contrasting historical demographies using three regions of the nuclear genome (appr. 3000 bps) in Europe. We aimed to draw inferences concerning phylogeography, demography and molecular evolutionary patterns of the species. *Sphagnum fimbriatum* and *S. squarrosum* are both monoecious but the former is presumably spreading in Europe. All three nuclear loci supported the presence of an Atlantic and Non-Atlantic clade of *S. fimbriatum* suggesting glacial survival of the species along the Atlantic coast of Europe. Contrarily, *S. squarrosum* haplotypes showed three clades but no geographic structure at all. Maximum likelihood, mismatch and Bayesian analyses supported a relatively recent demographic expansion of the Non-Atlantic clade of *S. fimbriatum*, whereas size of *S. squarrosum* populations has probably decreased in the past. Species wide molecular diversity of the two species was nearly the same with an excess of replacement mutations in *S. fimbriatum.* Similar level of molecular diversity of the species and excess of replacement mutations in *S. fimbriatum* compared to *S. squarrosum* can both be attributed to differences in outcrossing rate, demography and population history of the species. This study represents the first detailed European wide phylogeographic investigation on bryophytes and provides the first information how demography, population history and reproductive features shape molecular variability in haploid plants.

Keywords: *Sphagnum*, phylogeography, demography, multilocus, isolation with migration, mating system

Introduction

Several factors shape the spatial distribution and structure of genetic variation in plants. These include demography, population structure, mating system and historical effects caused by climatic or other factors. All of these are further enhanced by evolutionary forces such as mutation, migration, drift and selection.

Influence of environmental changes has governed genetic variability of species worldwide (Hewitt 1996). Distribution and amount of neutral genetic variability has been investigated in numerous species. In spite of individual differences according to species, results showed traces of a common historical event (Hewitt 2000; Taberlet *et al.* 1998) supposed to be the Quaternary glaciations, which influenced the worldwide distribution of the biota (Hewitt 1996; 1999).

The neutral mutation parameter is influenced by the mutation rate and effective population size. Mating system does not affect mutation rate but it influences the effective population size of genes. Inbreeding in diploid plants leads to high homozygosity within populations, which results in lower effective population size of genes (Charlesworth & Wright 2001; Charlesworth 2003). This latter is further reduced by the decreased effective recombination rate due to high homozygosity. Reduced effective population size increases the effect of genetic drift and leads to reduced molecular variability (Charlesworth 2003). In addition, because of the low effective recombination rate, hitchhiking events (like selective sweep or background selection) can extend their effect to a larger portion of the genome and thus further reduce molecular variability (Maynard Smith & Haigh 1974; Charlesworth *et al.* 1993). Consequently, within population variability of inbreeders should show less than half of the molecular variability of an outcrosser (Nordborg 2000; Nybom 2004).

Magnitude of species wide genetic variability in inbreeders and outbreeders might be very similar (Liu *et al.* 1998; Wright *et al.* 2003; Bakker *et al.* 2006), when migration is restricted among populations (Ingvarsson 2002; Panell & Charlesworth 2001). Weedy life form, usually associated with inbreeding is characterized by frequent extinction and recolonization cycles, which in turn might decrease the overall level of genetic diversity (Panell & Charlesworth 2001; Ingvarsson 2002; Charlesworth 2003).

Mating system influences efficacy of selection as well (Charlesworth & Wright 2001). Reduced effective population size of genes in an inbreeder relaxes the efficacy of selection (Charlesworth 2003). Slightly deleterious variants might easily become fixed in inbreeding populations resulting in an excess of nonsynonymous amino acid polymorphisms compared to neutral expectations (Savolainen *et al.* 2000; Bustamante *et al.* 2002; Wright *et al.* 2003).

Historical demography, population structure and selection also considerably influence the distribution and level of genetic diversity (Rosenberg & Nordborg 2002; Hein *et al.* 2004). Testing the selective neutrality of evolution involves testing deviation from the neutral coalescent (Wright & Gaut 2004). Models of neutral evolution assume constant historical population sizes and no population structure. Violations of these assumptions lead to significant deviations from neutrality (Tajima 1989; Wright & Gaut 2004). Investigating noncoding, neutrally evolving sequence portions does not overcome the problem of selection because they might be influenced by hitchhiking effects of linked loci (Maynard Smith & Haigh 1974). In theory, demography and selection result in the same pattern of molecular variation, e.g. rapid coalescent events following long branches on the coalescent tree (Hein *et al.* 2004). In the mismatch distribution, long branches lead to an excess of low frequency polymorphisms (Slatkin & Hudson 1991; Rodgers & Harpending 1992). Although selective forces and demography might produce the same pattern of polymorphism, selection affects the genome locally, whereas demography extends its effect on the whole genome. Consequently, investigating multiple, unlinked loci can help to separate the influence of selection and demography on nucleotide polymorphism (Wright & Gaut 2004; Nielsen 2005).

Bryophytes have a haploid dominant life cycle, which leads to differently acting evolutionary forces as in diploid seed plants (Stenøien & Såstad 1999; Stenøien & Flatberg 2000). Genome of haploid gametophores is directly exposed to selective forces and no sheltering of recessive alleles is possible (Wyatt *et al.* 1989). There seems to be selection against high mutation rates in bryophytes as well, otherwise slightly deleterious mutations would accumulate and result in a severe genetic load. This would lead to reduced molecular variation compared to seed plants (Stenøien & Såstad 2001). Contrarily, gene flow among bryophyte populations is expected to be more effective than

in seed plants because of the numerous spores, which can travel longer distances (Korpelainen *et al.* 2005; Sundberg 2005; Sundberg *et al.* 2006). Although most bryophytes are not dispersal limited at a large time scale, only a small portion of spores can establish in nature (Longton 1997; Sundberg & Rydin 2002; Sundberg 2002). This suggests that bryophytes are likely to be much more influenced by genetic drift than seed plants.

Peat moss gametophores are haploid, female and male gametangia might be produced on the same (bisexual gametophores) or on separated gametophores (unisexual gametophores). Bisexual gametophores can fertilize themselves (intragametophytic selfing) yielding totally homozygous progenies (Shaw 2000) or another gametophore established from spores originating from the same capsule (intergametophytic selfing). Fertilization usually takes place among gametophores situated close to each other but that not necessarily means inter- or intragametophytic selfing take place (Zielinski 1984; Cronberg 1996) and several mechanisms have developed in nature to avoid them (Wyatt & Anderson 1984). High levels of inbreeding can be prevented by frequent gene-flow among populations mediated by spores, scattered occurrence of slightly related genotypes and differences in timing of gametangia development (Cronberg 2002; Cronberg *et al.* 2003). These characteristics lead to different partitioning of molecular variability among and within populations of bryophytes compared to diploid seed plants (Van der Velde & Bijlsma 2003).

Peat mosses are an ancient plant group with a worldwide distribution (Daniels & Eddy 1985). They are a model system for population genetic research in bryophytes (Cronberg 1996; Cronberg & Natcheva 2002; Shaw *et al.* 2005; Flatberg *et al.* 2006). Phylogenetically closely related species pairs with similar distribution patterns but different life history characteristics (breeding system, demography) provide an opportunity to test the influence of species-specific features on the molecular evolution of genes.

Using chloroplast sequences, contrasting patterns of molecular variability and phylogeographic structure were found in the monoecious *S. fimbriatum* and *S. squarrosum* in Europe (Szövényi *et al.* 2006a). *S. fimbriatum* showed two geographically well delimitated lineages and low overall nucleotide diversity. In contrast, *S. squarrosum*

haplotypes showed higher nucleotide diversity and European lineages had no geographic affinities. This pattern was attributed to differences in life history characteristics and historical population sizes during the last glaciations.

In this study, we used sequences of three nuclear regions covering *appr.* 3000 bps of the genome to investigate population history and molecular population genetics of the two species in detail. Using three independent loci we intended to differentiate between selective and demographic forces. We were particularly interested in the following questions: 1. Do all nuclear loci support the same phylogeographic history? 2. Does multilocus analysis support assumed population expansion of *S. fimbriatum*? 3. Is there any signs of difference in the reproductive characteristic of the two species at the level of DNA polymorphism?

Materials and methods

Sampling and DNA extraction

S. fimbriatum and *S. squarrosum* plants were sampled by the authors, provided by colleagues or obtained from herbaria (Appendix 1). Plants were dried and stored in silica gel at 4 C°. DNA was extracted using the DNAeasy plant mini kit (Qiagen, Switzerland) following the manufacturers protocol. Three regions of the nuclear genome were used for sequencing. The region spanning the 5th-9'th exons of the GapC (glyceraldehyde-3-phosphate-dehydrogenase C) gene (appr. 1500 bps) was amplified using protocols and primers described in Szövényi *et al.* (2006a). A genomic region with unknown function (RAPDa sensu Shaw et al. 2003a) was also used. New primers were designed to improve amplification and sequencing of the fragment. As a third locus the ITS1-5.8S-ITS2 region (hereafter ITS) were amplified and sequenced. Yield of PCR reactions were improved by applying a semi-nested PCR, where 1 µl of the 10× diluted first round product was used as template in the second reaction. Primer sequences, annealing temperatures and extension times are given in Table 1. The PCR reaction contained the same amount of components as described in Szövényi et al. (2006b) except that additional 1%$^{v}/_{v}$ DMSO was used to amplify the ITS region. PCR reactions were run on Biometra (T1 or Tgradient, Whatman Biometra) and Techne ptc 412 (Barloworld Scientific Ltd.) thermocyclers. The PCR program was the following: 94 ºC denaturation for 4 min, then 94 ºC for 1 min, different annealing temperatures for 1 min and 72 ºC extension for 1-1.5 min (Table 1) with 35 cycles with a final extension step at 72 ºC for 7 min. The products were checked on 0.8 % agarose gels and cleaned using the GFX PCR and gel band purification kit (Amersham Biosciences, Switzerland). Approximately 10 ng product was sequenced in 10 µl cycle sequencing reaction with the Big dye v3.1 in an ABI prism 3100 (Applied Biosystems) genetic analyzer using either the original and/or different internal PCR primers (Table 1). Sequences were contiged and corrected if necessary with the Sequencher 4.5 software (Gene Code Corporation).

Table 1 Usage and sequences of primers used in this study. Names of newly designed primers are in bold. 1st and 2nd refer to the first and the second round of a semi nested PCR. In the second round PCR 1 µl 10x diluted first round product was used as template.

Region	Name	Usage	Sequence (5`-3`)	Design based on	Annealing T and length of extension
ITS	ITS1	amplification	TCCGTAGGTGAACCTGCGG	White et al. 1990	1st 50 °C, 60 s
	ITS4	amplification and sequencing	TCCTTCCGCTTATTGATATGC	Baum et al. 1998	2nd 54 °C, 60 s
	ITS1-int.	amplification and sequencing	CACACAGAGCGGTAAACCCTGC	*Sphagnum* sequences	
RAPDa	A-F	amplification	AACCAAGTGAATTTGGAATGC	Shaw et al. (2003a)	1st 56 °C, 90 s
	A-R	amplification and sequencing	AGGAGCGGAAGGCAAATG	Shaw et al. (2003a)	2nd 57 °C, 90 s
	RAPDa-forw.	amplification and sequencing	GATCCAGCCAAATCCACAAGATTCA	*Sphagnum* sequences	
	RAPDa-rev.	sequencing	CCTTYGACAAGGTTCGTGKTCTACTC	*Sphagnum* sequences	
	RAPDa-forw.int.	sequencing	TCCTCGATCCAGBAGATGGTAGA	*Sphagnum* sequences	
	AiR	sequencing	CAGAATGGCGAGCTTCCT	Shaw et al. (2003a)	
	AiF	sequencing	CAGCATTTTGGCTTTCCAAG	Shaw et al. (2003a)	

Data analysis

Sequence alignment, molecular diversity and recombination event estimates

Sequences were aligned using ClustalW (Thompson *et al.* 1994), checked by eye and adjusted if needed. All polymorphic sites were rechecked on the original electropherograms and corrected to avoid false base callings.

Haplotype networks were reconstructed using the TCS software (Clement *et al.* 2000). To describe the infra specific molecular diversity of the regions Watterson`s ($\theta_w$) (Hudson 1990) and Tajima`s (θ_π) (Tajima 1989) moment estimators of the neutral mutation parameter were calculated. In case of the GapC gene, levels of polymorphism were calculated for synonymous, nonsynonymous and silent sites as well.

All loci were tested for recombination using the four gamete test and the minimum number of recombination events during the history of the sample was calculated (R_m, Hudson & Kaplan 1985). For several following analyses (cf. below) sequences with incompatible nucleotide pairs were cut into recombination free blocks. All calculations were done using dnaSP v.4.10.4 (Rozas *et al.* 2003).

Testing deviation from neutrality

To test whether the sequenced regions evolved under neutral expectations, four different methods were applied. The multilocus HKA test (Hudson *et al.* 1987) was conducted using the HKA software (http://lifesci.rutgers.edu/~heylab). 10 000 coalescent simulations were run to generate the null distribution of χ^2 values. Neutral evolution of the GapC gene was also tested with the McDonald-Kreitman test (McDonald & Kreitman 1991). To test the neutraliy of each locus separately, Tajima's D (Tajima 1989) and Fay and Wu's H statistics (Fay & Wu 2000) were calculated for each of them. For each locus test statistics were calculated considering nucleotide substitutions alone and coding gaps as nucleotide substitutions as well.

Analysis of population growth

As a powerful descriptive statistic of population growth, Fu`s F_s (Fu 1997), was calculated using frequency distribution of alleles for each region separately. Populations

after a demographic expansion show a star-like phylogeny and a unimodal mismatch distribution (Slatkin & Hudson 1991), where the peak corresponds to the time of the expansion. The method of Schneider and Excoffier (1999) implemented in ARLEQUIN 3.0 (Excoffier *et al.* 2005) was used to fit observed and expected mismatch distributions and to estimate parameters of a pure demographic and a spatial expansion model.

As an alternative, maximum likelihood estimation of the exponential growth rate and historical θ was conducted. The software LAMARC 2.0 (Kuhner *et al.* 2005) was used to get estimates of growth rates for all European accessions of *S. squarrosum* and *S. fimbriatum* and for the Non-Atlantic clade of *S. fimbriatum*. 30 short chains (20 000 steps each) and 2 long chains (200 000 steps) with a sampling interval of 20 steps and a burn in period of 1000 and 10 000 respectively were applied.

The IM software (Hey & Nielsen 2004; Hey 2005) was used to fit the data to the Isolation with Migration model (IM). This model assumes two populations connected by migration, which are derived by the splitting of an ancient population. Populations were defined based on our a priory knowledge of chloroplast lineages using the parsimony networks of each species (Szövényi *et al.* 2006a). Runs were conducted for the combined dataset including all three nuclear loci sequenced. Initial runs with wide priors were used to delimit the plausible range of priors. Since indels of all three loci were informative, these have been coded as nucleotides. The infinite site mutation model of sequence evolution was applied and loci were cut into non-recombining blocks. In case of the RAPDa sequences of *S. squarrosum*, only the microsatellite region was used. Each run was introduced by a 100 000 steps long burn-in period. Multilocus analyses were run using 5–7 parallel chains under a linear heating scheme with a heating value of 0.05 – 0.005. Convergence of parameters and mixing of chains were followed by visual inspection of parameter trend lines and checking of ESS values. Analyses were run until the lowest ESS value reached minimum 200. In all analyses population demographic changes were allowed.

Results

Geographic distribution of haplotypes

In *S. fimbriatum*, The GapC and the RAPDa loci resolved the highest number of haplotypes (19 and 16 respectively), whereas the ITS region had much less resolution (9 haplotypes). All three loci supported a well-defined split of haplotypes into two groups (Fig 1). One lineage, further referred to as "Atlantic clade" occurs along the Atlantic coast of Spain, France and southern part of Britain. The rest of the accessions grouped into a clade extending from Southern France to Scandinavia ("Non-Atlantic clade") with one or two frequent and several rare haplotypes. In contrast to the ITS and RAPDa regions, the GapC gene gives further interpretable resolution within the Non-Atlantic clade. Plants from eastern part of Spain (haplotypes G and H, Fig 1) form a separate, geographically well-delimitated clade, also accessions from Austria and Hungary (haplotype E, except one occurrence in Scandinavia, Fig 1) show some geographic affinity.

In *S. squarrosum*, all three nuclear markers provided similar distributions of haplotypes, however, markers differed much more in their resolution than in *S. fimbriatum* (Fig 2). The ITS region contained only two singleton polymorphisms and one deletion, which divided the accessions into two haplotype groups with similar number of accessions. The GapC gene resolved 11 haplotypes, whereas in the RAPDa region no substitutions were found except two informative indels. In spite of the lack of substitutions, the RAPDa region contained a complete dinucleotide repeat, which showed considerable variability and resolved 20 haplotypes. The dinucleotide repeat was also present in *S. fimbriatum* sequences but showed no variability. In contrast to *S. fimbriatum*, *S. squarrosum* haplotypes show no clear geographic affinity. One of the ITS types tends to be more frequent in the south than in the north, but with a considerable admixture. RAPDa haplotypes show no clear geographic pattern either. Haplotypes of the GapC gene are also widely distributed and only two of them show any geographic grouping (D, G Fig 2).

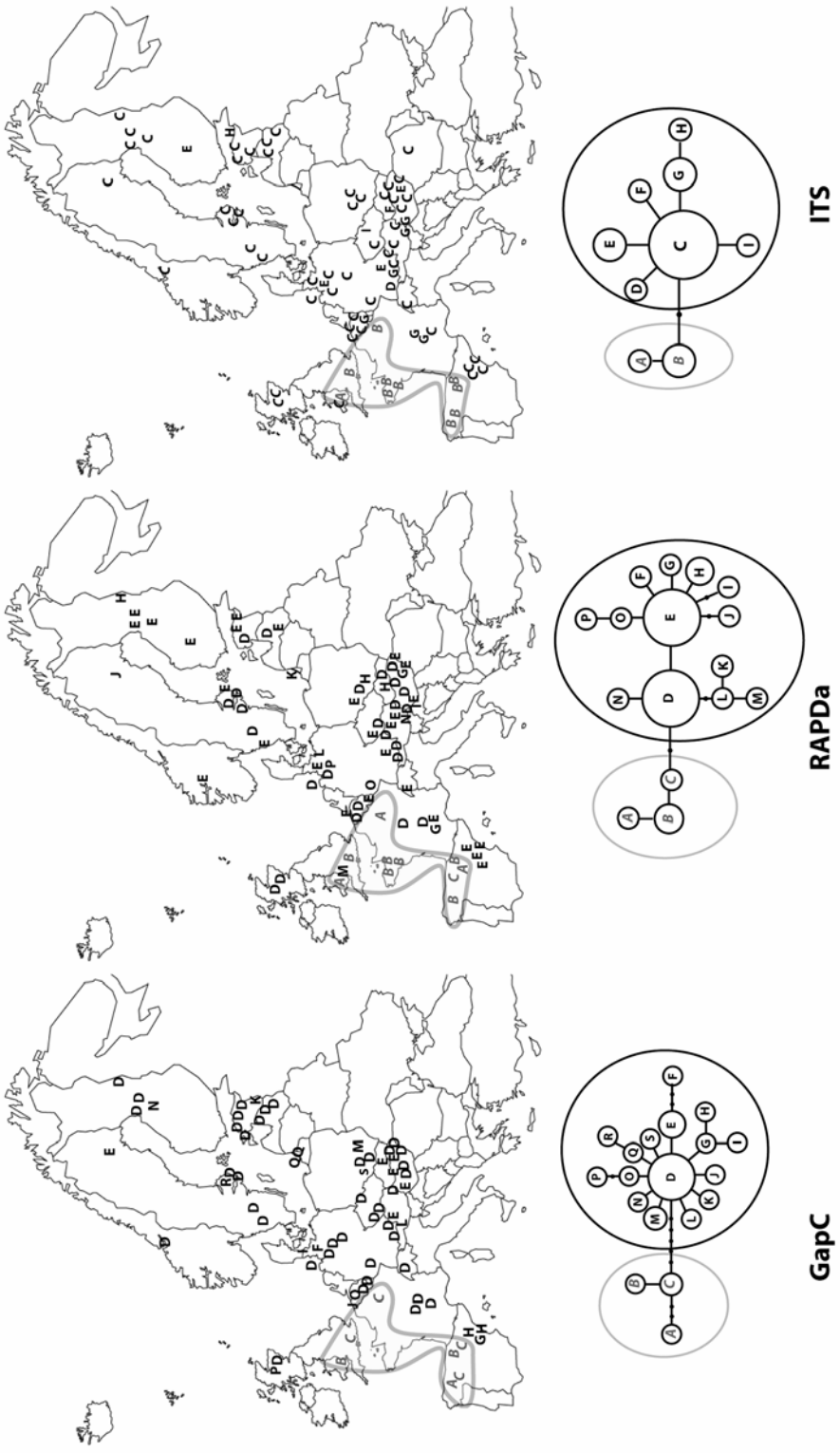


Fig 1 Geographic distribution of *S. fimbriatum* haplotypes and the corresponding maximum parsimony networks for each nuclear region investigated. In each haplotype network, letters in circles represent haplotypes found. Missing haplotypes are marked with black dots and lines connecting haplotypes denote one mutational change. The Atlantic and Non-Atlantic clades are shown within gray and black ellipses respectively.

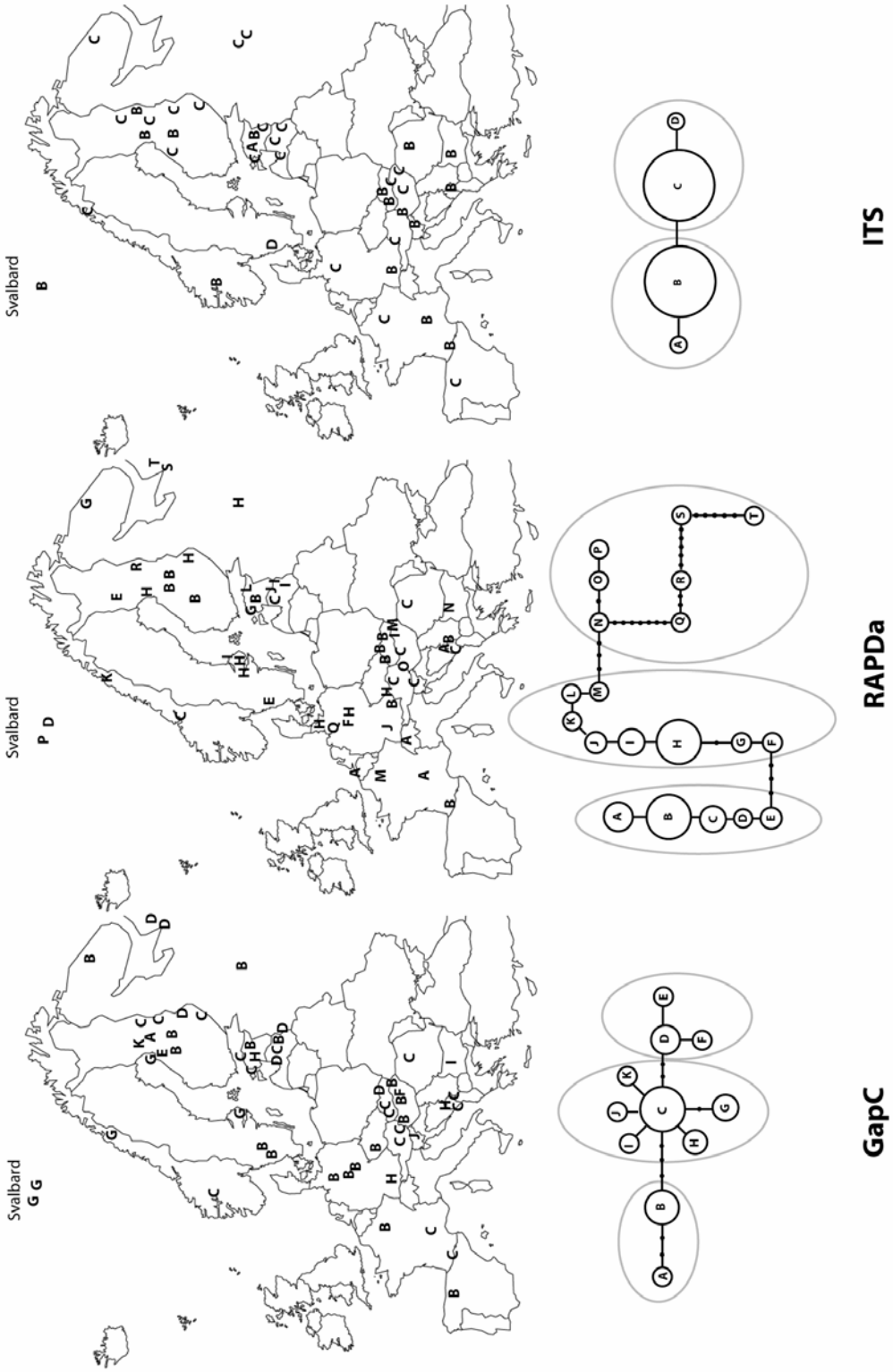


Fig 2 Geographic distribution of *S. squarrosum* haplotypes and the corresponding maximum parsimony networks for each nuclear region investigated. In each haplotype network, letters in circles represent haplotypes found. Missing haplotypes are marked with black dots and lines connecting haplotypes denote one mutational change. Well differentiated clades are shown within gray and black ellipses respectively.

Molecular polymorphism, tests of neutrality and recombination

Estimates of θ were lower in *S. squarrosum* for the ITS region but were similar for the GapC gene in both species (Table 2). The GapC gene of *S. fimbriatum* showed high diversity at synonymous sites whereas in *S. squarrosum* all mutations were found in noncoding regions.

The RAPDa region has been excluded from the multilocus HKA test because it showed no point mutations in *S. squarrosum* and only the microsatellite repeat was variable. No significant deviation from the neutral model was detected using the multilocus HKA test (p=0.48). Despite of the excess of non-synonymous substitutions in *S. fimbriatum* compared to *S. squarrosum* (Table 2), the McDonald-Kreitman test (GapC gene) was not significant either (p=0.50). In *S. squarrosum*, Tajima`s D and Fay and Wu`s H statistics were never significantly different from the neutral expectations.

Using all European accessions of *S. fimbriatum*, Tajima`s D was significant for all three nuclear loci investigated

Table 2 Number of sequences, aligned length, nucleotide diversity estimates and neutrality tests of the nuclear regions investigated. Tajima`s D was calculated taking nucleotide substitutions or nucleotide substitutions and gaps into account. θw total: Watterson`s estimator calculated using all sequences, θπ total: Tajima`s estimator calculated using all sequences, syn: synonymous sites, nonsyn: nonsynonymous sites, sil: silent sites.

	N	Aligned lenght in bps	$\theta_{w\ total}$	$\theta_{\pi\ total}$	$\theta_{w\ syn}$	$\theta_{w\ nonsyn}$	$\theta_{w\ sil}$	Tajima`s D (with gaps/without gaps)			Fay and Wu H
								Atlantic clade	Non-Atlantic clade	Total data set	Total data set
S. fimbriatum											
RAPDa	80	993	0,00103	0,00120	-	-	-	-0.6909/-1.1117	-1.7448*/0.4741	-1.3992*/0.3600	0,4044
ITS	90	614	0,00164	0,00077	-	-	-	-1.1117/-1.1117	-1.7460*/-1.6393*	-1.5950*/-1.1427	0,4155
GapC	76	1566	0,00236	0,00101	0,01124	0,00186	0,00267	-0.8764/-1.3272	-2.2321*/-1.9777*	-1.9074*/-1.8090*	-3,2821
S. squarrosum											
RAPDa	57	1033	-	-	-	-	-	-	-	-	-
ITS	60	668	0,00033	0,00005	-	-	-	-	-	-0.2501/-1.0857	0,0655
GapC	50	1561	0,00201	0,00202	0,00000	0,00000	0,00201	-	-	-0.4287/0.0279	-1,2114

θ_w total: Watterson`s estimator calculated using all accessions; $\theta_\pi$ total: Tajima`s estimator calculated using all accessions;
syn: synonymous sites; nonsyn: nonsynonymous sites; sil: silent sites; * : p≤0.05.

when including gaps (Table 2). However, it turned out to be non-significant when excluding gaps, except for the GapC gene. Fay and Wu`s H was never significant. Analysing accessions of the Non-Atlantic clade of *S. fimbriatum*, Tajima`s D was significantly negative except the RAPDa locus. The latter showed a significant negative value as well when gaps were included. The Atlantic clade showed no significant deviations from neutrality.

No signs of recombination were found in RAPDa and ITS data sets of both species. GapC sequences of *S. fimbriatum* showed one incompatible pair and an R_m value of 1. The two accessions causing the incompatibility were removed from the data set in further analyses. No signs of recombination were found in GapC sequences of *S. squarrosum*.

Population growth

Without gaps, Fu`s F_s statistic was only significant in the case of the GapC gene including all European accessions of *S. fimbriatum* (Table 3). In contrast, when gaps were used in the calculations, it turned out to be significantly different from a constant population in all three regions investigated. Analysing only the Non-Atlantic group of accessions, test statistic was significantly negative in all three

Table 3 Test of population growth using Fu`s F_s statistic for each nuclear regions investigated. Calculations were done including only nucleotide substitutions or nucleotide substitutions and gaps in the analysis.

	Fu`s F_s (with gaps/without gaps)		
	Atlantic	Non-Atlantic	Total data set
S. fimbriatum			
RAPDa	-0.5938/-0.3393	-8.4558*/0.7145	-8.0260*/0.7850
ITS	-0.3393/nc	-7.1366*/-5.2437*	-6.3052*/-1.7310
GapC	0.5410/0.8564	-15.4390*/-11.5356*	-10.7240*/-7.4648*
S. squarrosum			
RAPDa	-	-	-
ITS	-	-	-0.4350/-1.7750
GapC	-	-	-0.6240/1.2670

* : $p \leq 0.05$

genomic regions when including gaps. None of these statistics were significant using European accessions of *S. squarrosum* and accessions of the Atlantic clade of *S. fimbriatum*.

Due to low substitutional variability, ITS and RAPDa regions of both species had

Table 4 Test of population growth using mismatch distribution and maximum likelihood estimation. Mismatch analysis was done using all European accessions of both species. Note that estimations were done only for the GapC gene.

Mismatch analysis								
	Pure demographic expansion				Spatial expansion			
	p	Tau	θ_0	θ_1	p	Tau	θ_0	
S. fimbriatum	0,026	-	-	-	0,710	0.325 (0.210-4.710)	0.782 (0.000-1.652)	
S. squarrosum	0,143	4.976 (1.860-9.362)	0.000 (0.000-1.813)	6.509 (2.843-88.003)	0,257	4.195 (1.678-7.446)	0.092 (0.000-1.995)	

Tau: scaled time elapsed since the expansion; θ_0: theta of the population before expansion; θ_1: theta of the population after expansion; 90% confidence intervals are given in brackets.

Maximum likelihood estimates				
	Non-Atlantic clade		Europe	
	θ	g	θ	g
S. fimbriatum	0.01218 (0.00393-0.06008)	15000.00 (7831.62-32513.97)	0.00377 (0.00176-0.00678)	857.04 (-622.83-5891.93)
S. squarrosum	-	-	0.00242 (0.00113-0.00563)	233.58 (-1244.40-2401.97)

θ: maximum likelihood estimators of the neutral mutation parameter; g: the exponential growth parameter; 95% percentile intervals of values around the most likely estimates are shown in brackets.

to be excluded from the maximum likelihood estimation of exponential growth rate and historical theta. In *S. fimbriatum* estimation was made separately for all European accessions and for the Non-Atlantic clade. Analysis using all European accessions of *S. fimbriatum* showed about 4 times higher growth rates than in *S. squarrosum*, however, values were not significantly different from a shrinking, stable or expanding population (Table 4). Historical theta values were nearly the same for both species. Estimation provided two orders of magnitude higher growth rates and about one order of magnitude higher theta estimates for the Non-Atlantic clade of *S. fimbriatum* compared to *S. squarrosum*. Growth rate of *S. fimbriatum* in the Non-Atlantic clade was significantly different from zero, whereas that of *S. squarrosum* might come from a declining, expanding or stable population as well. Although likelihood surfaces were relatively flat, growth rate values of the two species turned out to be significantly different taking approximate confidence intervals into account.

In the mismatch distribution analysis, only the GapC sequences were used because of the low variability of the RAPDa and ITS regions (Table 4). The sudden demographic and the spatial expansion model both fit mismatch data of *S. squarrosum*. However, the fit was better using the spatial expansion model. The mismatch distribution of all European accessions of *S. fimbriatum* was significantly different from the pure demographic expansion model, but matched the spatial expansion model. Tau was at least 4 times larger in *S. squarrosum* than in *S. fimbriatum*. Theta estimated by the spatial expansion model was greater in *S. fimbriatum*, but this difference was not statistically significant. Estimation of parameters of sudden population growth or population expansion models failed for the Non-Atlantic group of *S. fimbriatum*, because the non-linear least squares algorithm failed to converge in ARLEQUIN.

Fitting the data to the IM model

The provided population genetic parameter estimates of the IM model are all scaled. In the absence of reliable mutation rate estimates of the regions investigated results were interpreted in a relative way (Fig 3). Current theta values of the Atlantic and Non-Atlantic clades of *S. fimbriatum* showed considerably different distributions and

point estimates. The Atlantic clade had lower theta than the Non-Atlantic clade and theta of the ancestral population was much lower than both of them. Estimation of theta for the Non-Atlantic clade was flat and converged but did not reach zero even after considerable extension of the prior. Scaled migration parameters were all very close to zero. Point estimate of the splitting parameter was high.

Sphagnum fimbriatum *Sphagnum squarrosum*

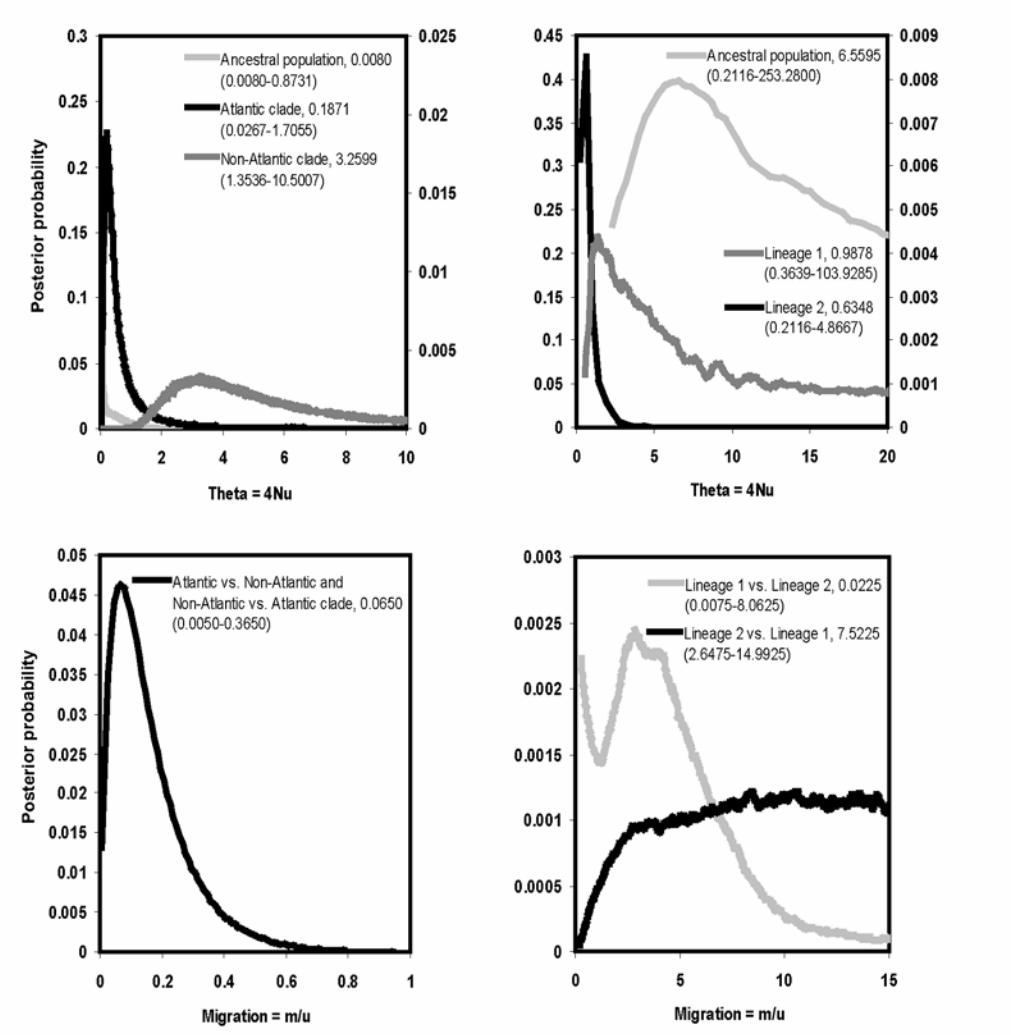


Fig 3 Posterior probability density functions of the parameters estimated using the Isolation with Migration model. In the figures point estimates of each parameter and the 90% highest posterior density intervals (in brackets) are given for each parameter. Note that parameter estimates are scaled and gaps were also used in the analyses. N: effective population size, m: rate of migration per gene per generation, u: mutation rate per gene per generation.

In contrast to *S. fimbriatum*, current thetas in *S. squarrosum* almost gave the same point estimates. Point estimate of the ancestral neutral mutation parameter was an order of magnitude higher than that of the current populations, but the posterior distribution was relatively flat. It converged to zero at higher theta values but did not reach zero even after extending the prior of theta considerably. Scaled migration parameters showed non zero values but were flat. Posterior probability distribution of the splitting parameter was flat and uninformative (not shown).

Discussion

Phylogeographic structure

Using chloroplast markers, considerable geographic structuring of haplotypes was detected in *S. fimbriatum*, in contrast to *S. squarrosum*, which showed no phylogeographic structure at all (Szövényi *et al.* 2006b). Results of the present analysis confirm these findings and support the formerly detected presence of an Atlantic and a Non-Atlantic lineage in *S. fimbriatum*. As proposed previously, this split would represent the effect of the Quaternary cold periods. Our current results evidently support a genome wide historical event, which influenced all loci in the same way regardless of genomic position.

Since there is no reliable calibration point in the history of the genus, it is not possible to date the split. In absence of dating at least two hypotheses are plausible. The first one assumes that Quaternary cold periods were long enough to allow accumulation of the observed number of mutations among the Atlantic and Non-Atlantic groups of populations. In addition, it presumes that current geographic position of populations roughly correspond to that of the historical ones. The Atlantic group might have survived along the Atlantic coast of Spain, France and South England, but location of refugia of the Non-Atlantic group remains ambiguous. Chloroplast markers suggested survival of this latter group on the eastern side of the Iberian Peninsula (Szövényi *et al.* 2006b). Better resolution obtained with nuclear markers rather suggests that Eastern Iberian populations represent recent colonization events. Using reconstructions of Gajewski *et al.* (2001) based on peat moss spores we hypothesize survival of the Non-Atlantic lineage along the border of Poland and Slovakia where extensive *Sphagnum* dominated habitats occurred in about 19 ky before present.

The second hypothesis assumes intercontinental migration between Europe and America. North American and European populations of several *Sphagnum* species appear to be closely related (Wyatt *et al.* 1992; Shaw & Schneider 1995; Såstad *et al.* 2000; Shaw *et al.* 2003c). Among others, long-range dispersal as a plausible explanation for the observed pattern has been suggested (Van Zanten 1978; Cronberg 2002; Sundberg *et al.*

2006) and the present data seem to support it. Assuming effective intercontinental migration, colonization from a divergent American population might account for the difference between the two lineages. However, testing this hypothesis need additional samples from America, which is the topic of our ongoing research. Although comparison with American haplotypes is not yet possible, fitting the IM model to the data set suggests that a very high proportion of the ancient population founded the Atlantic clade and the size of the ancient population was very small. These findings correspond to an intercontinental migration hypothesis including a strong founder effect during colonization. However, a combined scenario with separation of the two lineages following colonization from the New World might also be possible.

Unfortunately, our data set does not allow to clearly adjudge the plausibility of the two hypotheses mentioned above. However, several issues indicate that the Quaternary period might have played an important role in the observed genetic structure of the species. First, the patterns detected in this study were very similar to those found in other bryophyte species (Cronberg 1998; Van der Velde & Bijlsma 2003) and seed plants (Rendell & Ennos 2002; Hofmann *et al.* 2003). The second reason comes from fossil data and climatic reconstructions. Occurrence of relic *Sphagnum* species (Daniels & Eddy 1985), fossil remains of peat moss spores or leaves (Forment & Jovet-Ast 1950; Infante & Heras 1987) and climatic reconstructions also support survival of populations along the Atlantic coast of Europe (Gajewski *et al.* 2001; VanAndel & Tzedakis 1996). Based on our results and the above-mentioned reasons, we support the hypothesis assuming the separation of Atlantic and Non-Atlantic lineages by Quaternary glacial cycles and to reject the one presuming recent colonization from another continent as the cause of the split. However, colonization from out of Europe preceding divergence cannot be excluded.

Population demography

Selection influences each gene differently, whereas demographic processes affect the individual as a whole and thus influence all parts of the genome uniformly regardless of physical position (Wright and Gaut 2004). Hence, in an expanding population an

excess of low frequency variants is expected for all loci. This leads to negative Tajima`s D and Fu`s F_s statistics (Tajima 1989; Ramos-Onsins & Rosas 2002). These statistics all analyse the total or relative number of singleton mutations compared to the neutral coalescent. Their power depends on the total number of segregating sites in the data set and the time elapsed since the demographic event (Ramos-Onsins & Rosas 2002; Sano & Tachida 2005). Therefore it is not surprising that some of the statistics were not significant when excluding gaps as segregating sites from the analyses. This shows that substitution rates are relatively low compared to the time elapsed since the demographic event. Increasing the mutation rate with including gaps as mutations in the calculations improves power of the tests and extends the time frame in which those are applicable. Using these assumptions, all test statistics for all regions of *S. fimbriatum* were significantly different from the neutral expectations, whereas none of them provided significant deviations in *S. squarrosum*. This confirms a general excess of low frequency variants, in *S. fimbriatum* especially in the Non-Atlantic clade. Since the regions used are unlinked, we favour a population demographic expansion against a genome wide selection hypothesis in *S. fimbriatum*. This is further confirmed by the very similar patterns revealed in the study using chloroplast genes (Szövényi *et al.* 2006b). Sequencing errors are negligible because both species were analysed using the same protocols.

Mismatch distribution analysis revealed further details about the demographic history of the two species. Mismatch distribution of spatial expansion only slightly differs from that of pure demographic expansion, the latter showing a higher frequency of identical sequences (Excoffier 2004). In case of *S. fimbriatum*, data support the spatial expansion model, which is in concordance with the previous hypotheses about rapid spatial expansion of the species after the Quaternary glaciations. Data from *S. squarrosum* fit both scenarios, however, considerably better match the latter. This suggests that *S. squarrosum* probably went through spatial expansion as well. Tau value of *S. squarrosum* was approximately an order of magnitude higher than that of *S. fimbriatum*, although they were not significantly different. This might suggest that expansion of *S. fimbriatum* is more recent than that of *S. squarrosum*. Assuming that *S. fimbriatum* expanded immediately after the Last Glacial Maximum (LGM), the estimate

for *S. squarrosum* suggests expansion probably related to earlier geological or climatic events than the LGM. However, results of mismatch analyses need to be interpreted with caution. These analyses do not use genealogical information included in the data and testing the fit between observed and expected distributions and estimating approximate confidence intervals represent a complex problem as well. Estimation and confidence intervals of tau and θ are skewed and biased, especially in data sets with low resolution (Schneider & Excoffier 1999).

Maximum likelihood (LAMARC) and Bayesian analyses (IM) also supported rapid population expansion in *S. fimbriatum*. Results of both analyses showed no significant evidence for population growth in *S. squarrosum*. The IM analysis even supported a population decline since the separation of the two lineages of *S. squarrosum*.

Effect of outcrossing rate on polymorphism pattern

Mating system, in connection with evolutionary forces, can considerably alter the magnitude of species-wide and within population genetic diversity (Stenøien & Såstad 2001; Charlesworth 2003). *S. fimbriatum* has bisexual gametophytes and frequently produces sporophytes, which indicates frequent intra- and intergametophytic selfing. *S. squarrosum* has also bisexual gametophytes but functionally unisexual gametophores are frequently found in nature (Daniels & Eddy 1985), suggesting a lower level of inbreeding.

Effect of mating system on the total nucleotide diversity of the species is mainly influenced by population structure, migration rate among populations and fluctuation of population size. Overall diversity values for the ITS region were usually an order of magnitude higher in *S. fimbriatum* than in *S. squarrosum*, however values were very low. Moreover in *S. fimbriatum* a considerable portion of the diversity was due to the genetic differentiation between Atlantic and Non-Atlantic clades. In contrast, GapC sequences of the species showed nearly the same nucleotide diversity values. This is an unexpected result because the pioneer characteristic of *S. fimbriatum* (Sundberg *et al.* 2006), even combined with a parallel historical population size reduction of *S. squarrosum* implies lower level of molecular polymorphism in the first species. Peat moss populations are

likely to function as a metapopulation (Söderström & Herben 1997; Wakeley & Alicar 2001). Theoretical expectations show that genetic diversity will be lost rapidly in metapopulations of pioneer species with frequent extinction/recolonization processes regardless of outcrossing rate or among population migration rates (Panell & Charlesworth 2001; Ingvarsson 2002). Consequently, similar level of nucleotide diversity in *S. fimbriatum* and *S. squarrosum* can hardly be explained by differences in metapopulation processes.

Although magnitude of molecular polymorphism was similar in both species its pattern differed considerably showing an excess of polymorphism in coding regions of *S. fimbriatum*. Recent increase of effective population size might result in a significant excess of slightly deleterious replacement mutations (Eyre-Walker 2002). Weakened selection due to the reduced effective population size can lead to the accumulation of slightly deleterious mutations as well and thus to elevated nucleotide diversity in selfing plants (Charlesworth & Wright 2001). The McDonald-Kreitman test showed no significant deviations from neutral expectations for the GapC gene. However, exons of this gene contained replacement mutations in *S. fimbriatum* whereas in *S. squarrosum* all mutations were restricted to noncoding regions. Replacement mutations in *S. fimbriatum* (Lys-Arg, Ile-Val, Lys-Gln) caused no change in hydrophoby and almost no change in pI values and thus are likely to be slightly deleterious. High level of molecular diversity in coding regions of *S. fimbriatum* compared to *S. squarrosum* implies, that higher inbreeding rates in *S. fimbriatum* might have also played a role in the accumulation of replacement mutations. Assuming similar rates of inbreeding in both species would have certainly led to at least some mutations in coding regions of *S. squarrosum*. This is further suggests that there seems to be considerable difference in outcrossing rates of *S. fimbriatum* and *S. squarrosum* although both are morphologically monoecious. These findings are in line with the relaxed selective expectation and recent population growth, which contribute significantly to the nucleotide diversity of *S. fimbriatum*. However, outcrossing rates of the species in nature are not known yet and must be determined experimentally (Cronberg 1996).

Historical isolation of populations might also have contributed to the nucleotide diversity of *S. fimbriatum*. The species is easily dispersed by spores to longer distances

(Szurdoki & Ódor 2005; Sundberg *et al.* 2006), however, chloroplast and nuclear markers both resolved distinct Atlantic and Non-Atlantic clades in Europe. All markers indicate either restricted migration or establishment potential measured on the time scale of the markers' mutation rates. *S. fimbriatum* produces considerable amount of spores, which can germinate on a wide range of substrates (especially on substrates with low phosphate availability) compared to other *Sphagnum* species (Sundberg & Rydin 2002; Sundberg *et al.* 2006). Based on these characteristics of the species, the Non-Atlantic clade of *S. fimbriatum* might have rapidly recolonized available soil surfaces after the last glaciation. In absence of further space, the Atlantic clade might have been unable to follow this lineage and remained restricted to the Atlantic coast of Europe. Similar patterns of leading edge colonization have been observed in several tree species in Europe (Demesure *et al.* 1995; Heuertz *et al.* 2004).

It has been assumed that the two frequent European chloroplast haplotype lineages of *S. squarrosum* correspond to populations that were formerly isolated (Szövényi *et al.* 2006b). The IM analysis suggested a considerable population size reduction in both European lineages of *S. squarrosum*. It is hard to date this event, but we suppose that peat exploiting activities during the last centuries might have caused it. This population size reduction might have further decreased genetic diversity of the species compared to *S. fimbriatum*.

Overall nucleotide diversity of the three regions was low in both species compared to estimations in seed plants (reviewed in Wright & Gaut 2004), which might be a general trend in bryophyte populations. Due to haploid gametophores no sheltering of recessive alleles exists in bryophytes and the haploid genome of each individual is directly exposed to selection (Wyatt *et al.* 1992). This, coupled with the presence of deleterious mutations, considerably reduces the amount of genetic diversity. Mutation rate also influences genetic variability. Selection against high mutation rates should exist in bryophytes; otherwise the presence of slightly deleterious mutations will lead to a severe mutational load (Stenøien & Såstad 2001). Our results show that slightly deleterious mutations are present but are removed very efficiently. The theory of selection directly acting on the gametophores and low mutation rates are supported by the data set presented here. Assuming the same mutation rates in both species, selection is so

efficient, that no replacement mutations were found in the GapC gene of *S. squarrosum* at all. In contrast to our results, high levels of isozyme diversity were found within populations of several bryophytes (Wyatt *et al.* 1992; Stenøien & Såstad 1999). This has partly been interpreted as a result of local adaptation (Shaw 2000). Although local adaptation might be important in several species, our results do not support this hypothesis, since only one of the three replacement mutations in GapC gene resulted in charge change. Hence these mutations probably do not influence enzyme functions significantly.

References

Bakker EG, Stahl EA, Toomajian C et al. (2006) Distribution of genetic variation within and among local populations of *Arabidopsis thaliana* over its species range. *Molecular Ecology*, **15**, 1405-1418.

Baum DA, Small RL, Wendel JF (1998) Biogeography and floral evolution of baobabs (*Adansonia*, Bombacaceae) as inferred from multiple data sets. *Systematic Biology*, **47**, 181-207.

Bustamante CD, Nielsen R, Sawyer S et al. (2002) The cost of inbreeding in *Arabidopsis*. *Nature*, **416**, 531-534.

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, **134**, 1289-1303.

Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. P*hilosophical Transactions of the Royal Society of London series B*, **358**, 1051-1070.

Charlesworth D, Wright SI (2001) Breeding system and genome evolution. *Current Opinion in Genetics and Development*, **11**, 685-690.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1660.

Cronberg N (1996) Clonal structure and fertility in a sympatric population of the peat mosses *Sphagnum rubellum* and *Sphagnum capillifolium*. *Canadian Journal of Botany*, **74**, 1375-1385.

Cronberg N (1998) Population structure and interspecific differentiation of the peat moss sister species *Sphagnum rubellum* and *S. capillifolium* (Sphagnaceae) in northern Europe. *Plant Systematics and Evolution*, **209**, 139-158.

Cronberg N (2002) Colonization dynamics of the clonal moss *Hylocomium splendens* on islands in a Baltic land uplift area: reproduction, genet distribution and genetic variation. *Journal of Ecology*, **90**, 925-935.

Cronberg N, Andersson K, Wyatt R, Odrzykoski IJ (2003) Clonal distribution, fertility and sex ratios of the moss *Plagiomnium affine* (Bland.) T.Kop. in forests of contrasting age. *Journal of Bryology*, **25**, 155-162.

Cronberg N, Natcheva R (2002). Hybridization between the peat mosses, *Sphagnum capillifolium* and *S. quinquefarium* (Sphagnaceae, Bryophyta) as inferred by morphological characters and isozyme markers. *Plant Systematics and Evolution*, **234**, 53-70.

Daniels RE, Eddy A (1985) Handbook of European *Sphagna*. Institute of Terrestrial Ecology, Huntingdon.

Demesure B, Comps B, Petit RJ (1995) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515-2520.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, **13**, 853-864.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.

Eyre-Walker A (2002) Changing effective population size and the McDonald-Kreitman test. *Genetics*, **162**, 2017-2024.

Fay JC, Wu CI (2000) Hitchhiking under positive darwinian selection. *Genetics*, **155**, 1405-1413.

Flatberg KI, Thingsgaard K, Såstad SM (2006) Interploidal gene flow and introgression in bryophytes. *Sphagnum girgensohnii* x S. *russowii*, a case of spontaneous neotriploidy. *Journal of Bryology*, **28**, 27-37.

Forment P, Jovet-Ast S (1950) *Sphagnum imbricatum* Russ., subfossil en France dans une tourbe des environs de Calais. *Revue Bryologique et Lichénologique*, **19**, 125-126.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Gajewski K, Viau A, Sawada M, Atkinson D, Wilson S (2001) *Sphagnum* peatland distribution in North America and Eurasia during the past 21,000 years. *Global Biogeochemical Cycles*, **15**, 297-310.

Hein J, Schierup MH, Wiuf C (2004) Gene Genealogies, Variation and Evolution. A Primer in Coalescent Theory. Oxford University Press, Oxford.

Heuertz M, Fineschi S, Anzidei M, et al. (2004) Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. *Molecular Ecology*, **13**, 3437- 3452.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.

Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Hey J (2005) On the Number of New World Founders: A Population Genetic Portrait of the Peopling of the Americas. *PLoS Biol*, **3**:e193.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis. Genetics*, **167**, 747-760.

Hofmann MH, Glass AS, Tomiuk J, et al. (2003) Analysis of molecular data of *Arabidopsis thaliana* (L.) Heyn. (Brassicaceae) with geographical information systems (GIS). *Molecular Ecology*, **12**, 1007-1019.

Hudson RR. 1990. Gene genealogies and the coalescent process. In: Harvey PH, Partridge L, eds. *Oxford Surveys in Evolutionary Biology. Vol. 7.* Oxford University Press, USA: New York, 1-44.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147-164.

Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153-159.

Infante M, Heras P (1987) *Sphagnum imbricatum* Hornsch. ex Russ., semifosil, en la turbera de Saldropo-Brazar (Ceanuri, Vizcaya). *Estudios Instituto Alaves de la Naturaleza*, **2**, 201-203.

Ingvarsson PK (2002) A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution*, **56**, 2368-2373.

Korpelainen H, Pohjamo M, Laaka-Lindberg S (2005) How efficiently does bryophyte dispersal lead to gene flow? *Journal of the Hattory Botanical Laboratory*, **97**, 195-207.

Kuhner MK, Yamato J, Beerli P, et al. (2005) LAMARC v 2.0. University of Washington, http://evolution.gs.washington.edu/lamarc.html 2005.

Liu F, Zhang L, Charlesworth D (1998) Genetic diversity in *Leavenworthia* populations with different inbreeding levels. *Proceedings of the Royal Society of London B*, **265**, 293-301.

Longton RE (1997) Reproductive Biology and Life-History Strategies. In: *Advances in Bryology* **6** (ed. Longton RE), pp. 65-103. Science Publishers, Stuttgart.

Maynard Smith J, Haigh J (1974) The hitchhiking effect of a favourable gene. *Genetic Research Cambridge*, **219**, 1114-1116.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the ADH locus in *Drosophila*. *Nature*, **351**, 652-654.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197-218.

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics*, **154**, 923-929.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143-1155.

Panell JR, Charlesworth B (2001) Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. *Evolution*, **53**, 664-676.

Ramos-Onsins SE, Rosas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092-2100.

Rendell S, Ennos RA (2002) Chloroplast diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69-78.

Rodgers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552-569.

Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, **3**, 380-390.

Rozas J, Sanches-Delbarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496-2497.

Sano A, Tachida H (2005) Gene genealogy and properties of test statistics of neutrality under population growth. *Genetics*, **169**, 1687-1697.

Såstad SM, Flatberg IK, Hanssen L (2000) Origin, taxonomy and population structure of the allopolyploid peat moss *Sphagnum majus*. *Plant Systematics and Evolution*, **225**, 73-84.

Savolainen O, Langley CH, Lazzaro BP, Freville H (2000) Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. *Molecular Biology and Evolution*, **17**, 645-655.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079-1089.

Shaw AJ (2000) Population ecology, population genetics, and microevolution. In: *Bryophyte Biology* (eds. Shaw AJ, Goffinet B), pp. 369-403. Cambridge University Press, Cambridge.

Shaw AJ, Cox CJ, Boles SB (2003a) Polarity of peatmoss evolution: who says mosses have no roots? *American Journal of Botany*, **90**, 1777-1787.

Shaw AJ, Cox CJ, Boles SB (2003b) Global patterns in peatmoss biodiversity. *Molecular Ecology*, **12**, 2553-2570.

Shaw AJ, Melosik I, Cox CJ, Boles SB (2005) Divergent and reticulate evolution in closely related species of *Sphagnum* section Subsecunda. *The Bryologist*, **108**, 363-376.

Shaw AJ, Schneider RE (1995) Genetic biogeography of the rare „copper moss“ *Mielichhoferia elongata* (Bryaceae). *American Journal of Botany*, **82**, 8-17.

Shaw AJ, Werner O, Ros RM (2003c) Intercontinental mediterranean disjunct mosses: morphological and molecular patterns. *American Journal of Botany*, **90**, 540-550.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555-562.

Söderström L, Herben T (1997) Dynamics of Bryophyte Metapopulations. In: *Advances in Bryology* **6** (ed. Longton RE), pp. 205-241. Science Publishers, Stuttgart.

Stenøien HK, Flatberg KI (2000) Genetic variability in the rare Norwegian peat moss *Sphagnum troendelagicum*. *The Bryologist*, **103**, 794-801.

Stenøien HK, Såstad SM (1999) Genetic structure in three haploid peat mosses (*Sphagnum*). *Heredity*, **82**, 391-400.

Stenøien HK, Såstad SM (2001) Genetic variability in bryophytes: does mating system really matter? *Journal of Bryology*, **23**, 313-319.

Sundberg S (2002) Sporophyte production and spore dispersal phenology in *Sphagnum*: the importance of summer moisture and patch characteristics. *Canadian Journal of Botany*, **80**, 543-556.

Sundberg S (2005) Sporophyte size positively influences short-range spore dispersal in *Sphagnum*, but what happens further away? *Oikos*, **108**, 115-124.

Sundberg S, Hansson J, Rydin H (2006) Colonisation of *Sphagnum* on land uplift islands in the Baltic Sea: time, area, distance and life history. *Journal of Biogeography*, **33**, 1479-1491.

Sundberg S, Rydin H (2002) Habitat requirements for establishment of *Sphagnum* from spores. *Journal of Ecology*, **90**, 268-278.

Szövényi P, Hock Zs, Urmi E, Schneller JJ (2006a) New primers for amplifying the GapC gene of bryophytes and its utility in infraspecific phylogenies in the genus *Sphagnum*. *Lindbergia* (in press).

Szövényi P, Hock Zs, Urmi E, Schneller JJ (2006b) Contrasting phylogeographic patterns in *Sphagnum fimbriatum* and *S. squarrosum* (Bryophyta, Sphagnopsida) in Europe. *New Phytologist*, doi: 10.1111/j.1469-8137.2006.01870.x.

Szurdoki E, Ódor P (2005) Distribution and expansion of *Sphagnum fimbriatum* Wils. in Hungary. *Lindbergia*, **29**, 136-142.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-

specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680.

Van der Velde M, Bijlsma R (2003) Phylogeography of five *Polytrichum* species within Europe. *Biological Journal of the Linnean Society*, **78**, 203-213.

VanAndel TH, Tzedakis PC (1996) Paleolithic landscapes of Europe and environs, 150,000-25,000 years ago: An overview. *Quaternary Science Reviews*, **15**, 481-500.

Van Zanten, BO (1978) Experimental studies on trans-oceanic long-range dispersal of moss spores in the southern hemisphere. *Journal of the Hattori Botanical Laboratory*, **44**, 455-482.

Wakeley J, Alicar N (2001) Gene genealogies in a metapopulation. *Genetics*, **159**, 893-905.

White TJ, Burns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: In: A Guide to Methods and Applications (eds. Innis M, Gelfand D, Sninsky J, White T), pp. 315-322. Academic Press, San Diego, California.

Wright SI, Gaut BS (2004) Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution*, **22**, 506-519.

Wright SI, Lauga B, Charlesworth D (2003) Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. *Molecular Ecology*, **12**, 1247-1263.

Wyatt R, Anderson LE (1984) Breeding systems of bryophytes. In: *The Experimental Biology of Bryophytes* (eds. Dyer AF, Duckett JG), pp. 39-64. Academic Press, London.

Wyatt R, Odrzykoski I, Stoneburner J (1992) Isozyme evidence of reticulate evolution in mosses: *Plagiomnium medium* is an allopolyploid of *P. ellipticum* x *P. insigne*. *Sytematic Botany*, **17**, 532-550.

Wyatt R, Stoneburner A, Odrzykoski IJ (1989) Bryophyte isozymes: Systematic and evolutionary implications. In: *Isozymes in Plant Biology* (eds. Soltis DE, Soltis P), pp. 221-240. Dioscorides Press, Portland.

Zielinski, R (1984) Electrophoretic evidence of cross-fertilization in the monoecious *Pellia epiphylla* N=9. *Journal of the Hattory Botanical Laboratory*, **56**, 255-262.

Acknowledgements

We thank Adam Hölzer, Alain Vanderpoorten, Alessia Guggisberg, Ana Seneca, Andrea Fodor, Andrej Martinčič, Anna Ganeva, Baiba Bambe, Beáta Papp, Beatriz Itten, Blanka and Jonathan Shaw, Christian Dolnik, Claes Persson, Edwin Urmi, Erzsébet Szurdoki, Harald Zechmeister, Heike Hofmann, Ildikó László, Irina Goldberg, José Pujos, Jürgen Brandt, Kai Vellak, Liudmila Abramova, Luca Miserere, Markus Preussing, Marta Infante, Michael Ignatov, Monika Koperski, Nele Ingerpuu, Patxi Heras, Péter Ódor, Philippe de Zuttere, Pierre Goubet, Risto Virtanen, Robert Krisai, Ron Porley, Rudolf Soltés, Sam Bosanquet, Sebastian Sundberg, Sigurd Såstad, Sture Westerberg, Zoltán Illyés and Zoltán Tóth for collecting material and/or for field assistance. This research was financially supported by a Young Academics Grant of the University of Zurich (Forschungskredit der Universität Zürich) and by a Travel Grant of the Swiss National Academy of Science.

Appendix 1 Detailed list of the *Sphagnum fimbriatum* and *S. squarrosum* accessions analysed for the three regions in this study. Letters refer to haplotypes of Figure 1 and 2. na: not analysed.

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype			GenBank accession number		
					GapC	RAPDa	ITS	GapC	RAPDa	ITS
Sphagnum fimbriatum	1	Scotland, Forest of Alyth	400 m	56° 30' N 03° 20' W	P	D	C	SEQUENCES SENT TO GENBANK BUT ACCESSION NUMBERS NOT YET RECEIVED		
Sphagnum fimbriatum	2	Scotland, Trinafour	300 m	56° 45' N 03° 55' W	D	D	C			
Sphagnum fimbriatum	8	Hungary, Alsó-Erdö I	330 m	47° 24' N 16° 33' E	M	D	C			
Sphagnum fimbriatum	9	Hungary, Büdöskút-Árpádforrás	435 m	47° 22' N 16° 28' E	D	E	G			
Sphagnum fimbriatum	10	Hungary, Büdöskút-Árpádforrás	435 m	47° 22' N 16° 28' E	D	I	C			
Sphagnum fimbriatum	14	Norway, Sor-Trondelag	10 m	63° 42' N 08° 46' E	D	E	C			
Sphagnum fimbriatum	56	Bohemia, Trebon	-	48° 46' N 14° 42' E	D	na	C			
Sphagnum fimbriatum	63	Bohemia, Stare Jezero	-	48° 46' N 14° 42' E	D	D	C			
Sphagnum fimbriatum	65	Bohemia, Krkonoše Mts.	1020 m	50° 45' N 15° 33' E	D	na	I			
Sphagnum fimbriatum	67	Bohemia, Chlum u trebone	-	49° 27' N 13° 35' E	D	E	C			
Sphagnum fimbriatum	69	Hungary, Velencei-tó	-	47° 11' N 18° 33' E	D	D	C			
Sphagnum fimbriatum	70	Hungary, Szigetcsép Csupics-sziget	-	47° 15' N 18° 59' E	E	D	C			
Sphagnum fimbriatum	72	Hungary, Dorog	150 m	47° 43' N 18° 45' E	E	na	F			
Sphagnum fimbriatum	75	Hungary, Monostori-tó	345 m	46° 54' N 17° 35' E	D	D	C			
Sphagnum fimbriatum	79	Austria, Elixhausen	560 m	47° 53' N 13° 01' E	L	E	C			
Sphagnum fimbriatum	84	Upper Austria, Tarsdorf	480 m	48° 05' N 12° 51' E	D	D	C			
Sphagnum fimbriatum	93	Germany, Durchenbergried	430 m	47° 45' N 08° 58' E	D	na	D			
Sphagnum fimbriatum	100	Germany, Hüven	10 m	52° 46' N 07° 34' E	D	E	C			
Sphagnum fimbriatum	102	Germany, Butterloch	200 m	51° 35' N 10° 18' E	D	na	C			
Sphagnum fimbriatum	108	Sweden, Mulön	-	65° 37' N 22° 17' E	D	D	C			
Sphagnum fimbriatum	109	Sweden, Bensbyn	-	65° 38' N 22° 13' E	E	J	C			
Sphagnum fimbriatum	110	Germany, Neunkirchen	10 m	52° 46' N 08° 44' E	D	D	E			
Sphagnum fimbriatum	114	Germany, Hamberger Moor	15 m	53° 16' N 08° 51' E	F	P	C			
Sphagnum fimbriatum	118	France, Troncais forest	280 m	49° 52' N 04° 38' E	D	D	G			
Sphagnum fimbriatum	120	Spain, Candin, Lamela	850 m	42° 57' N 06° 41' W	A	B	B			
Sphagnum fimbriatum	121	Spain, Onate, Sierra de Elguea-Urkilla	1045 m	42° 57' N 02° 24' W	B	B	B			
Sphagnum fimbriatum	122	Spain, Orihuela del Tremedal	1675 m	40° 32' N 01° 38' W	na	E	C			
Sphagnum fimbriatum	132	Belgium, Nonceveux	100 m	50° 26' N 05° 43' E	O	D	C			
Sphagnum fimbriatum	137	Latvia, sample 4	-	56° 40' N 25° 55' W	D	E	C			
Sphagnum fimbriatum	138	Latvia, sample 3	-	56° 40' N 25° 55' W	D	D	C			

Continued

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype			GenBank accession number		
					GapC	RAPDa	ITS	GapC	RAPDa	ITS
Sphagnum fimbriatum	139	Latvia, sample 2	-	56° 37' N 26° 20' E	D	F	C			
Sphagnum fimbriatum	145	Germany, Rheinland-Pfalz	120 m	49° 24' N 07° 42' E	D	O	C			
Sphagnum fimbriatum	148	Hungary, Templom-tó	250 m	47° 01' N 16° 49' E	E	D	C			
Sphagnum fimbriatum	149	Hungary, Fias-tó	250 m	59° 38' N 19° 17' E	na	E	C			
Sphagnum fimbriatum	150	Hungary, Füzes-tó	250 m	47° 02' N 16° 48' E	na	na	C			
Sphagnum fimbriatum	152	Hungary, Kócse-tó	250 m	47° 01' N 16° 49' E	E	N	G			
Sphagnum fimbriatum	153	France, Walleis-Auemberg	-	50° 14' N 03° 41' E	C	A	B			
Sphagnum fimbriatum	166	Finland, Ruovesi	-	62° 04' N 24° 21' E	na	E	E			
Sphagnum fimbriatum	169	Finland, Ylikiiminki	-	64° 53' N 26° 07' E	D	E	C			
Sphagnum fimbriatum	170	Finland, Oulu	-	64° 34' N 25° 33' E	D	E	C			
Sphagnum fimbriatum	171	Finland, Kuusamo	-	65° 48' N 29° 50' E	D	H	C			
Sphagnum fimbriatum	172	Finland, Piippola	-	64° 13' N 25° 48' E	N	E	C			
Sphagnum fimbriatum	173	Belgium, Province of Liège	-	50° 12' N 05° 26' E	D	E	G			
Sphagnum fimbriatum	174	Belgium, Province of Liège	300 m	50° 14' N 05° 27' E	D	D	C			
Sphagnum fimbriatum	178	Slovakia, Poprad Basin	680 m	49° 03' N 20° 17' E	E	D	C			
Sphagnum fimbriatum	234	Sweden, Lilla Idskär (island)	3 m	59° 38' N 19° 17' E	R	D	C			
Sphagnum fimbriatum	235	Sweden, Storön (island)	2 m	59° 27' N 19° 30' E	na	D	C			
Sphagnum fimbriatum	236	Sweden, Rödskäret (island)	5 m	59° 36' N 19° 27' W	D	E	C			
Sphagnum fimbriatum	237	Sweden, Norra Skrakskär (island)	5 m	59° 35' N 19° 20' E	D	D	na			
Sphagnum fimbriatum	251	Austria, Kneisselmoor	-	47° 45' N 13° 01' E	E	D	C			
Sphagnum fimbriatum	252	Austria, Zehmemoos	-	47° 59' N 12° 55' E	D	D	E			
Sphagnum fimbriatum	253	Switzerland, Les Ponts-de-Martel	1000 m	46° 58' N 06° 43' E	D	E	C			
Sphagnum fimbriatum	254	Austria, Wengermoor	-	47° 55' N 13° 10' E	D	E	G			
Sphagnum fimbriatum	257	Hungary, Nyíres-tó	150 m	47° 49' N 22° 25' E	D	E	C			
Sphagnum fimbriatum	258	Hungary, Nyíres-tó	150 m	47° 49' N 22° 25' E	D	E	C			
Sphagnum fimbriatum	260	France, locality 2005/2	-	45° 51' N 02° 10' E	D	E	C			
Sphagnum fimbriatum	261	Spain, Barranco Larreakorta	700 m	43° 00' N 02° 49' W	C	C	B			
Sphagnum fimbriatum	262	Spain, Peña la Gallina	1620 m	40° 32' N 01° 42' W	G	E	C			
Sphagnum fimbriatum	266	France, locality 2005/4	-	45° 49' N 01° 59' E	D	D	G			
Sphagnum fimbriatum	267	Spain, Los Ojos	1310 m	40° 32' N 01° 38' W	G	E	C			
Sphagnum fimbriatum	268	Spain, Los Ojos	1310 m	40° 32' N 01° 38' W	H	E	C			

SEQUENCES SENT TO GENBANK BUT ACCESSION NUMBERS NOT YET RECEIVED

Continued

Species	Number of populations	Location	m a.s.l.	Longitude/latitude	Haplotype			GenBank accession number		
					GapC	RAPDa	ITS	GapC	RAPDa	ITS
Sphagnum fimbriatum	269	Spain, Candin, Lamela	861 m	42° 57' N 06° 41' W	C	B	B			
Sphagnum fimbriatum	297	Poland, Babezyna Dolina	-	49° 46' N 19° 03' E	na	na	C			
Sphagnum fimbriatum	298	Poland, Silesian lowland	-	51° 06' N 18° 22' E	M	D	C			
Sphagnum fimbriatum	299	Poland, Pagory jaworznickie hills	-	50° 07' N 19° 13' E	D	na	na			
Sphagnum fimbriatum	299	Poland, Pagory jaworznickie hills	-	50° 07' N 19° 13' E	na	E	C			
Sphagnum fimbriatum	300	Hungary, Regéc	300 m	48° 26' N 21° 26' E	na	na	C			
Sphagnum fimbriatum	301	Rumania, Hargita	-	47° 43' N 25° 11' E	na	na	C			
Sphagnum fimbriatum	302	Poland, Beskid Makowski Mountains	425 m	49° 40' N 19° 45' E	D	na	E			
Sphagnum fimbriatum	303	Hungary, Lókosár	-	48° 28' N 20° 30' E	D	G	E			
Sphagnum fimbriatum	304	Hungary, Nagymohos	290 m	48° 15' N 20° 13' E	E	D	C			
Sphagnum fimbriatum	305	Hungary, Springs of Tegda Valley	300 m	48° 23' N 21° 34' E	na	D	C			
Sphagnum fimbriatum	306	Hungary, Kismohos	300 m	48° 15' N 20° 13' E	na	H	C			
Sphagnum fimbriatum	307	Poland, Beskid Slaski Mts.	860 m	49° 38' N 19° 09' E	S	H	G			
Sphagnum fimbriatum	340	Germany, Spiekeroog (island)	2 m	53° 46' N 07° 41' E	D	D	E			
Sphagnum fimbriatum	342	Russia, Rybachy	5 m	55° 09' N 20° 49' E	Q	K	C			
Sphagnum fimbriatum	343	Germany, Jagen	12 m	54° 23' N 09° 32' E	I	L	C			
Sphagnum fimbriatum	344	Germany, Jagen	12 m	54° 21' N 10° 03' E	D	E	na			
Sphagnum fimbriatum	350	UK, Cors Farlais, Carmarthenshire	300m	51° 54' N 04° 06' W	C	A	A			
Sphagnum fimbriatum	351	UK, Rhos Rydd, Cardiganshire	-	04° 44' N 52° 03' W	D	M	C			
Sphagnum fimbriatum	357	Sweden, Göteborg, Halland	10 m	57° 31' N 12° 10' E	D	E	C			
Sphagnum fimbriatum	366	France, Louargat	-	48° 35' N 03° 37' W	na	B	B			
Sphagnum fimbriatum	367	France, Belle Isle-en-Terre	-	48° 28' N 03° 37' W	na	B	B			
Sphagnum fimbriatum	368	France, Bretagne	-	48° 13' N 03° 52' W	na	B	na			
Sphagnum fimbriatum	368	France, Louargat	-	48° 35' N 03° 37' W	na	na	B			
Sphagnum fimbriatum	372	Estonia, Nätsi	-	58° 30' N 24° 04' E	D	E	H			
Sphagnum fimbriatum	373	Estonia, Tartu	-	58° 23' N 27° 06' E	D	D	C			
Sphagnum fimbriatum	374	Estonia, Tartu	-	58° 23' N 27° 06' E	D	D	na			
Sphagnum fimbriatum	374	Estonia, Tartu	-	58° 23' N 27° 06' E	na	na	C			
Sphagnum fimbriatum	375	Estonia, Rapla county	-	59° 02' N 24° 29' E	K	E	C			
Sphagnum fimbriatum	403	Britain, Berkshire	-	51° 13' N 01° 13' W	C	B	B			
Sphagnum fimbriatum	409	Belgium, Stekene	-	51° 11' N 03° 59' W	J	E	C			
Sphagnum fimbriatum	410	Belgium, Matagne la Grande	-	50° 06' N 04° 38' E	na	na	C			

SEQUENCES SENT TO GENBANK BUT ACCESSION NUMBERS NOT YET RECEIVED

Species	Number of populations	Locality	m a.s.l	Longitude/Latitude	Haplotype			GenBank accession number		
					GapC	RAPDa	ITS	GapC	RAPDa	ITS
Sphagnum squarrosum	16	Norway, Sor-Trondelag	1 m	63° 42' N 08° 46' E	na	na	C	SEQUENCES SENT TO GENBANK BUT ACCESSION NUMBERS NOT YET RECEIVED		
Sphagnum squarrosum	17	Norway, Svalbard, Nordenskiöld Land	-	77° 55' N 14° 38' E	G	P	na			
Sphagnum squarrosum	18	Norway, Sogn og Fjordane	345 m	62° 00' N 10° 00' E	C	C	B			
Sphagnum squarrosum	19	Norway, Aure	20 m	63° 15' N 08° 31' E	na	E	B			
Sphagnum squarrosum	30	Pyrenees, Lac de Oredon	1900 m	42° 49' N 00° 10' E	C	B	B			
Sphagnum squarrosum	40	Switzerland, Eriz, Rotmoos	-	46° 54' N 07° 36' E	na	A	B			
Sphagnum squarrosum	81.1	Austria, Lungau	1260 m	48° 06' N 12° 52' E	na	B	B			
Sphagnum squarrosum	85.1	Austria, Filzmoos	480 m	48° 06' N 12° 53' E	B	H	C			
Sphagnum squarrosum	96	Germany, Wasenmoos	460 m	47° 41' N 09° 35' E	H	J	B			
Sphagnum squarrosum	101	Germany, Hüven	8 m	52° 46' N 07° 34' E	B	Q	C			
Sphagnum squarrosum	103	Germany, Butterloch	200 m	51° 35' N 10° 18' E	B	F	C			
Sphagnum squarrosum	105	Bulgaria, Vitosha mountains	1750 m	40° 22' N 23° 22' E	I	na	B			
Sphagnum squarrosum	107	Sweden, Mulön	-	65° 37' N 22° 17' E	B	N	na			
Sphagnum squarrosum	112	Germany, Göttingen	5 m	51° 34' N 10° 08' E	B	H	C			
Sphagnum squarrosum	134	Russia, Moscow	-	55° 30' N 37° 30' E	B	H	C			
Sphagnum squarrosum	135	Latvia, locality 7.	-	56° 40' N 25° 55' E	D	J	C			
Sphagnum squarrosum	136	Latvia, locality 8.	-	56° 37' N 26° 20' E	B	I	C			
Sphagnum squarrosum	140	Latvia, locality 1.	-	56° 40' N 25° 55' E	D	I	C			
Sphagnum squarrosum	141	Latvia, locality 5.	-	56° 37' N 26° 20' E	C	C	B			
Sphagnum squarrosum	143	Estonia, Nigula	10 m	24° 40' N 58° 01' E	C	G	C			
Sphagnum squarrosum	147	Hungary, Velencei-tó	100 m	47° 12' N 18° 33' E	B	C	C			
Sphagnum squarrosum	151	Hungary, Köcse-tó	250 m	47° 01' N 16° 49' E	B	O	na			
Sphagnum squarrosum	154	France, Walleis-Auemberg	-	50° 14' N 03° 41' E	B	M	C			
Sphagnum squarrosum	155	Finland, Hyrynsalmi	-	64° 19' N 29° 00' E	C	B	B			
Sphagnum squarrosum	156	Finland, Vörtsila	-	61° 56' N 29° 42' E	C	H	C			
Sphagnum squarrosum	157	Norway, Svalbard,Van Mijenfjorden	-	77° 49' N 16° 10' E	G	D	C			
Sphagnum squarrosum	158	Finland, Oulu	10 m	64° 58' N 25° 56' E	A	H	C			
Sphagnum squarrosum	159	Finland, Piipola	25 m	64° 22' N 25° 40' E	D	J	C			
Sphagnum squarrosum	160	Finland, Vieremä	-	63° 19' N 26° 32' E	B	B	B			
Sphagnum squarrosum	161	Finland, Hailuoto	10 m	65° 05' N 24° 46' E	G	R	C			
Sphagnum squarrosum	175	Slovakia, High Tatra mountains	1335 m	20° 03' N 49° 07' E	C	B	B			
Sphagnum squarrosum	176	Slovakia, High Tatra mountains	1080 m	49° 12' N 20° 16' E	C	B	na			
Sphagnum squarrosum	180	Slovakia, Poprad Basin	680 m	49° 03' N 20° 17' E	na	B	B			

Continued

Species	Number of populations	Locality	m a.s.l	Longitude/Latitude	Haplotype			GenBank accession number		
					GapC	RAPDa	ITS	GapC	RAPDa	ITS
Sphagnum squarrosum	239	Sweden, Långmossen	0 m	59° 58' N 17° 18' E	G	H	C	SEQUENCES SENT TO GENBANK BUT ACCESSION NUMBERS NOT YET RECEIVED		
Sphagnum squarrosum	240	Sweden, Norra Skrakskär (island)	3 m	59° 35' N 19° 20' E	na	I	C			
Sphagnum squarrosum	241	Sweden, Rödskäret (island)	2 m	59° 36' N 19° 27' E	na	H	C			
Sphagnum squarrosum	243	Sweden, Storön (island)	2 m	55° 11 N 20° 51' E	na	na	C			
Sphagnum squarrosum	264	Spain, locality 2005/9.	640 m	43° 02' N 02° 42' W	B	na	C			
Sphagnum squarrosum	265	France, locality 2005/3.	1030 m	45° 57' N 03° 41' E	C	A	B			
Sphagnum squarrosum	289	Austria, Seemoos	1670 m	47° 10' N 13° 47' E	C	C	B			
Sphagnum squarrosum	292	Romania, Kelemen havasok	800 m	47° 14' N 25° 16' E	C	C	B			
Sphagnum squarrosum	321	Poland, Oswiecim basin	-	50° 01' N 19° 06' E	na	na	C			
Sphagnum squarrosum	322	Hungary, Kismohos	300 m	48° 15' N 20° 13' E	F	M	C			
Sphagnum squarrosum	323	Hungary, Tegda	440 m	48° 25' N 21° 29' E	B	I	C			
Sphagnum squarrosum	324	Poland, Silesian upland	740 m	51° 11' N 18° 38' E	na	na	B			
Sphagnum squarrosum	325	Poland, Maly Mts.	-	50° 51' N 14° 26' E	na	na	B			
Sphagnum squarrosum	326	Poland, Rybnicki Plateau	-	50° 08' N 18° 27' E	na	na	B			
Sphagnum squarrosum	333.1	Montenegro, Durmitor	1320 m	42° 83' N 18° 23' E	C	C	na			
Sphagnum squarrosum	333.2	Montenegro, Durmitor	1320 m	42° 83' N 18° 23' E	C	C	B			
Sphagnum squarrosum	334	Serbia, Golija 03/105/6	1480 m	43° 20' N 20° 15' E	C	H	B			
Sphagnum squarrosum	335	Serbia, Golija	1480 m	43° 20' N 20° 15' E	na	B	B			
Sphagnum squarrosum	339	Norway, Troms	0 m	69° 47' N 21° 02' E	G	K	C			
Sphagnum squarrosum	341	Russia, Rybachy	2 m	55° 11' N 20° 51' E	na	na	C			
Sphagnum squarrosum	344	Germany, Jagen	12 m	54° 23' N 09° 31' E	A	na	C			
Sphagnum squarrosum	354	Russia, Bolshojslovetski Island	-	64° 42' N 39° 44' E	D	na	na			
Sphagnum squarrosum	354	Russia, Arckhangelsk region	-	64° 42' N 39° 44' E	na	T	C			
Sphagnum squarrosum	355	Russia, Arckhangelsk region	-	64° 42' N 39° 44' E	D	S	C			
Sphagnum squarrosum	356	Russia, Kola Peninsula	-	68° 05' N 39° 50' E	B	G	C			
Sphagnum squarrosum	360	Göteborg, Rammsjödal	10 m	57° 31' N 12° 10' E	B	E	D			
Sphagnum squarrosum	364	France, Spezet	-	48° 11' N 03° 42' W	na	na	C			
Sphagnum squarrosum	370	Estonia, Nigula	-	58° 23' N 27° 06' E	H	B	A			
Sphagnum squarrosum	371	Estonia, Tartu	-	58° 23' N 27° 06' E	B	L	C			
Sphagnum squarrosum	383	Slovenia, Šijec	1200 m	46° 20' N 13° 59' E	J	C	B			
Sphagnum squarrosum	393	Finland, Kuopio	-	62° 52' N 27° 41' E	K	B	B			
Sphagnum squarrosum	394	Finland, Liminka	-	64° 49' N 25° 23' E	E	E	C			
Sphagnum squarrosum	413	Belgium, Lux, Bellevaux	-	49° 50' N 05° 13' E	na	A	B			
Sphagnum squarrosum	414	Belgium, Slekene		51° 11' N 03° 59' W	na	F	C			

General summary

The three chapters of this thesis describe our investigations on the European phylogeography and historical demography of two peat mosses, using molecular markers. Besides giving a detailed description of the phylogeography of the two species we also test the most plausible hypotheses which may have led to the current geographic distribution of genetic variability of the two species in Europe.

In the first chapter, we investigated the European phylogeography of *Sphagnum fimbriatum* and *S. squarrosum*. This work constitutes, to the best of our knowledge, the first descriptive phylogeographic study on bryophytes. The two taxa have similar breeding systems, distributions and ecological requirements, but possibly different demographic histories, with a recent demographic expansion in *S. fimbriatum*. We demonstrate that phylogeographic structure of the two species is remarkably different. Geographically well-delimitated, slightly overlapping clades were found in *S. fimbriatum*, in contrast, *S. squarrosum* showed genetically differentiated but geographically well-mixed clades in Europe. Results were independent of the applied molecular markers (chloroplast or nuclear markers). Analyses of genetic data suggest that *S. fimbriatum* recolonized Europe very rapidly after the Last Glacial Maximum and that ability for expansion is an inherent characteristic of the species. It is hypothesized that this recent expansion of *S. fimbriatum* in Europe is a response to environmental/climate change. In addition, genetic data suggest that *S. fimbriatum*, a species less tolerant to desiccation compared to *S. squarrosum*, survived the last glaciations along the Atlantic coast of Europe and experienced a severe bottleneck. In contrast, *S. squarrosum* may have survived in several scattered refugia experiencing only a moderate bottleneck. Chloroplast markers proved to be useful to describe the phylogeographic structure of the two species but yielded insufficient amount of variability to statistically test contrasting hypotheses.

To obtain sufficient resolution for statistical tests, primers amplifying more variable parts of the peat moss genome have been developed and tested. In the second chapter, we designed new primers to amplify a single-copy gene of the bryophyte genome, the GapC gene. The structure of the GapC gene in peat mosses turned out to be

very similar to that of other green plants, except one additional intron disrupting exon nine. We carried out a preliminary screen for variability using some European accessions of *S. fimbriatum*. This analysis proved that the GapC gene provides more variable characters than the chloroplast markers of the previous study, and could be particularly useful in reconstructing infraspecific phylogenies and testing phylogeographic hypotheses in bryophytes. This preliminary analysis also confirmed the differentiation of the Atlantic and Non-Atlantic clades of *S. fimbriatum*.

In the third chapter, multilocus analyses of sequence data were conducted using coalescent, Bayesian and maximum likelihood estimations, to test whether genetic data support the hypothesis of a population expansion in *S. fimbriatum* and a more stable population size in *S. squarrosum*. All markers supported the presence of an Atlantic and a Non-Atlantic clade in *S. fimbriatum*, suggesting glacial survival of the species along the Atlantic coast of Europe. Conversely, *S. squarrosum* haplotypes showed three clades but no geographic structure at all. Maximum likelihood, mismatch and Bayesian analyses supported a relatively recent demographic expansion of the Non-Atlantic clade of *S. fimbriatum*, whereas sizes of *S. squarrosum* populations have probably decreased in the past. The species-wide molecular diversities of the two species were very similar, with an excess of replacement mutations in *S. fimbriatum*. These findings can be both attributed to differences in outcrossing rate, demography and population history between the two species.

It should be mentioned here that detailed population genetic analysis of several European populations have been conducted using AFLP markers in both species but due to the restricted amount of time, results could not be included in this thesis.

Beyond the remarkably different phylogeographic structure and demographic history of the species, investigations provided two main unexpected observations that need to be highlighted.

1.) The strong geographic affinity and only slightly overlapping structure of clades in *S. fimbriatum* compared to the frequent admixture of the *S. squarrosum* haplotypes.

2.) The extremely different pattern of polymorphisms found in the two species with an excess of replacement and synonymous mutations in *S fimbriatum* compared to *S. squarrosum*.

To explain these unexpected observations, differences in a wide range of biological aspects of the species need to be implicated. Such contrasting and unexpected results raise general questions concerning the rates of migration and admixture between populations and lineages, the molecular evolution of functionally important genes, the difference in outcrossing rates among species with the same mating system, and the worldwide demographic history of the species. The purpose of future research is to understand the biological basis for the phylogeographic structure of the two species. Population-level analyses will be carried out, using worldwide population samples and new genetic markers (such as microsatellites or/and SNPs).

Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Phylogeographie und der historischen Demographie zweier Torfmoosarten in Europa basierend auf molekularen Methoden. Das Ziel war einerseits ein umfassendes Bild der phylogeographischen Struktur der Arten in Europa zu geben und anderseits die Prozesse aufzuklären, die dieser Struktur zu Grunde liegen.

Im ersten Kapitel, wird die phylogeographische Struktur von *Sphagnum fimbriatum* und *S. squarrosum* in Europa dargestellt. Die Verbreitung der zwei Arten in Europa ist sehr ähnlich, sie kommen auch oft zusammen vor und beide sind monözisch. Ihre demographische Geschichte ist jedoch verschieden, indem *S. fimbriatum* eine rasche Ausbreitung seiner europäischen Populationen zeigt. Die phylogeographische Struktur der beiden Arten unterscheidet sich deutlich. Genetisch differenzierte Gruppen von *S. fimbriatum* haben eine geographisch beschränkte Verbreitung und zeigen nur geringe Überlappung. Demgegenüber zeigen die Haplotypen der genetisch differenzierten Gruppen von *S. squarrosum* eine geographische Durchmischung. Auch die genetischen Daten weisen darauf hin, dass *S. fimbriatum* Europa nach den letzten Eiszeiten schnell besiedelt hat. Es ist anzunehmen, dass die gegenwärtige Ausbreitung der Art in Europa mit dem Klimawandel zusammenhängt. Nach den vorliegenden Daten hat *S. fimbriatum* – eine Art mit geringer Austrocknungsresistenz – die Eiszeiten in kleinen Populationen entlang der atlantischen Küste Europas überdauert. *S. squarrosum* dagegen überlebte in mehreren zerstreuten Populationen näher am Eisrand und wurde deswegen vom Flaschenhalseffekt weniger beeinflusst. Obwohl Sequenzdaten der Chloroplastenmarker zur Aufklärung der phylogeographischen Struktur von Nutzen waren, ist ihre Variabilität zu gering, um unterschiedliche Hypothesen statistisch nachprüfen zu können.

Um eine statistische Auswertung zu ermöglichen, sind neue Primer für ein Single-Copy Gen des Kerngenoms (GapC Gen) entwickelt worden, und die Struktur und genetische Variabilität dieses Gens in Torfmoosen wurde geklärt. Dabei hat es sich herausgestellt, dass die Struktur des Gens jener der übrigen Landpflanzen gleicht mit der Ausnahme, dass das neunte Exon ein zusätzliches Intron enthält. Untersuchungen auf der Ebene der Sequenz-Variabilität haben ergeben, dass das GapC-Gen mehr Variabilität

aufweist als Chloroplastenmarker und deswegen für phylogeographische Untersuchungen sehr geeignet ist. Die vorläufigen Untersuchungen über die Variabilität des GapC-Genes haben die Resultate, die mit Hilfe der Chloroplastenmarker gewonnen wurden, bestätigt, nämlich, dass die genetische Differenzierung der atlantischen und kontinentalen Gruppen von *S. fimbriatum*. Diese Ergebnisse sind im zweiten Kapitel dargestellt

Das dritte Kapitel beschreibt eine eingehende Untersuchung mittels mehrerer Gene des Kerngenoms. Mit Hilfe von *Bayes-Statistik* und der c*oalescent theorie* wurde geprüft, ob europäische Populationen von *S. fimbriatum* ein rasches demographisches Wachstum aufweisen. Die genetischen Marker haben gezeigt, dass die antlantischen und kontinentalen Gruppen von *S. fimbriatum* genetisch differenziert sind, und die eiszeitlichen Refugien der Art entlang der atlantischen Küste von Europa lagen. Europäische Populationen von *S. squarrosum* gliedern sich auch in genetisch differenzierte Gruppen, aber diese kommen weitgehend gemischt vor und zeigen kaum geographische Präferenzen. *Maximum Likelihood*-, *Bayesian*- und *Mismatch*-Analysen zeigten, dass die Populationsgrösse der kontinentalen Gruppe von *S. fimbriatum* innerhalb kurzer Zeit sehr stark angewachsen ist. Dagegen hat die Populationsgrösse von *S. squarrosum* in der Vergangenheit wahrscheinlich leicht abgenommen. Das Ausmass der molekularen Diversität beider Arten ist ähnlich, aber *S. fimbriatum* zeigt mehr Punktmutationen als *S. squarrosum*. Die ähnliche molekulare Variabilität der beiden Arten zeigt, dass diese nicht nur hinsichtlich ihrer Demographien sondern wahrscheinlich auch in ihren historischen Prozessen und Selbsbefruchtungsraten verschieden sind.

Unsere Untersuchungen haben nicht nur sehr unterschiedliche phylogeographische Strukturen und demographische Abläufe in Europäischen Populationen von *S. fimbriatum* und *S. squarrosum* aufgedeckt, sondern auch zwei unerwartete Resultate betreffend geographische Verbreitung und molekulare Variabilität der Arten geliefert:

1.) Scharfe geographische Trennung der genetisch differenzierten Gruppen von *S. fimbriatum* im Gegensatz zu gemischten Vorkommen der genetisch differenzierten „clades“ von *S. squarrosum* in Europa.

2.) Das unterschiedliche Muster der molekularen Variabilität der zwei Arten, mit mehr Punktmutationen bei *S. fimbriatum*.

Diese Unterschiede lassen sich nur schwer ausschliesslich mit demographischen und/oder historischen Gründen erklären. Deswegen sollten noch weitere Faktoren in Betracht gezogen werden. Es ist anzunehmen, dass wichtige Unterschiede in der Biologie der beiden Arten vorkommen. Unsere unerwarteten Resultate haben uns zu weiteren Fragen geführt betreffend Migration Vermischung von genetisch differenzierten Populationen, molekulare Evolution der funktionalen Gene, Unterschiede in der Selbstbefruchtungsrate zwischen Populationen und weltweite demographische ‚Geschichte' der Arten. Zukünftige Untersuchungen sollten sich auf die möglichen Ursachen für diese unerwarteten Unterschiede konzentrieren. Um diese Ziele zu erreichen, sollen möglichst viele Proben von mehreren Populationen aus der ganzen Welt untersucht und neue molekulare Marker (z.B. Mikrosatelliten) entwickelt werden.

Curriculum vitae

Personalien

Name und Vorname:	Szövényi, Péter István Antal
Geboren am:	13. Juni 1975 in Budapest, Ungarn
Heimatort:	Budapest, Ungarn

Ausbildung

Mittelschule:	II. Rákóczi Ferencz Gymnasium, Budapest. Abschluss 1993, Matura
Hochschulstudium:	Eötvös Loránd Universität, Fakultät für Lehrerausbildung, Biologie und Chemie Lehrer, 1993 bis 1997 Eötvös Loránd Universität, Wissenschaftliche Fakultät, Fachbiologe, 1997 bis 2002
Diplomfach:	Evolutionsbiologie, Ökologie und Systematik
Diplomarbeit:	„Phorophyte preferences of epiphytic bryophytes in a stream valley: field studies and experimental investigations“. Lehrstuhl für Pflanzentaxonomie und Ökologie, Eötvös Loránd Universität, Wisseschaftliche Fakultät, Budapest, Ungarn Dr. Zoltán Tóth. 2002
Doktorat:	Universität Zürich, Institut für Systematische Botanik, Dr. Edwin Urmi und Dr. Jakob Schneller.